

South Africa's new landscape

Despite progress towards racial integration, proposals to rationalize by merging universities are being attacked as threats to the identity of some institutions. There is another way forward that deserves to be considered.

It is ironic that current plans to restructure South Africa's higher-education system — including the merger of many of the country's historically black institutions with historically white ones — are being met with so much resistance from the very institutions that serve to benefit most from the changes. But South Africa's education minister, Kader Asmal, has made it clear that, after many years of procrastination on this issue, the country should no longer be held to ransom by vested interests in the face of declining registrations at many of its universities.

Most of South Africa's historically white institutions have made significant progress in becoming nonracial, to the extent that they now have more black students than the country's historically black ones (see page 377). The challenge is to find the best way of extending this process to upgrading the historically black universities and technikons (polytechnics), most of which were set up during the apartheid era to cater for specific ethnic groups.

The plan currently under consideration by the South African cabinet, drawn up by the National Working Group on higher education, is to merge all but seven of the country's 21 universities and 15 technikons into 21 institutions. Twelve of the proposed 14 new institutions would involve mergers between historically black and historically white ones; the other two would be between historically black institutions.

This proposal might be workable if effective new management procedures are put in place, and could certainly promote broader access to the better resources of the historically white institutions. But it is questionable whether the proposed four new 'comprehensive institutions', which would offer both university and technikon programmes, could really work within the context of a system that retains a distinction between the two.

Two years ago, another higher-education government taskforce

proposed an alternative set of reform plans. It suggested dividing the country's universities and technikons into three categories: those that could award degrees up to doctoral level in all fields, those that could award only up to masters degrees in certain fields, and those that could award only undergraduate degrees.

Historically black institutions feared that they might fall into the last category, and their objections killed the proposals. But as problems with the latest plans are being aired, the previous model is once again being debated — if not by the cabinet, then at least within the universities themselves. The alternative model would at least allow institutions to retain their own identity — something they seem to be fearful of losing.

In addition, it would have the advantage of retaining undergraduate universities, at least, within reasonable access of the rural communities that several of them serve. Staff at these institutions could develop collaborative research programmes with established research universities, in a similar manner to the way that college staff do in the United States. If this were to happen, then researchers at these universities need not be denied the opportunity to continue with research, even if their universities were designated as undergraduate colleges. Many workable collaborations already exist between historically black and historically white universities in South Africa — both for teaching and research purposes — so in many instances there is a sound basis on which to expand.

Whichever model is adopted, the rationalization of the South African higher-education system will have implications for funding agencies, both within and outside the country. There is little point in research funds being directed at institutions that lack a research culture and are unlikely to acquire one. A new dispensation should precipitate increased international funds for research. It should also ensure that funds are better deployed. ■

Nature and the developing world

Nature's publishers are making their journals freely available to the world's poorest countries.

At the beginning of this year, a report from the World Health Organization (WHO) highlighted the potential benefits for the developed world in giving more aid to developing countries. The scale of increased aid that was recommended was big, according to economist Jeffrey Sachs and his committee, but so were the potential benefits (see *Nature* 415, 1; 2002).

Making scientific and medical journals available cheaply or at zero cost was just one component of Sachs' wish-list. To its credit, the WHO was able to begin to fulfil that hope with a scheme that it launched in January, ponderously entitled the Health Internetwork Access to Research Initiative (HINARI). In this scheme, publishers collaborate with the WHO's Health Internetwork (see www.healthinternetwork.net), which acts as an online gateway and validator for libraries, 490 of which have been identified by the WHO. Publishers make their information freely available from their own resources, without payment from third parties.

At present, free access through HINARI is provided to most of the 106 countries that fall below a threshold of having an annual gross national product per head of US\$1,000. From next year it is intended that journals will also be available at greatly reduced prices to the 39 countries whose gross national product per head falls between \$1,000 and \$3,000.

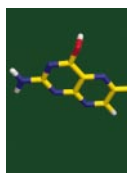
From 24 May, *Nature* and all other journals published by the Nature Publishing Group will become available through the HINARI scheme. That, coupled with the SciDev.Net free-access website that was spun off from *Nature* last year (see www.scidev.net), is one step towards the goals outlined by Sachs' committee.

In practice, the HINARI scheme has encountered some teething difficulties but, according to those in charge, it now works effectively. The biggest obstacle to access, they say, comes from "poor and extremely expensive connectivity". Telecommunication providers and governments, please note. ■





Sound bites
Study tracks
sophisticated views
on transgenics
p368



Name game
Chemists seek a
route from structure
to a proper name
p369



DNA clampdown
UK commission calls
for bar on illicit use
of DNA samples
p370



Killer question
US and Canada
clash over orphaned
orca
p372

Bell Labs launches inquiry into allegations of data duplication

Geoff Brumfiel, Washington

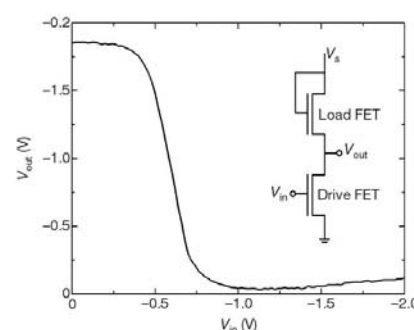
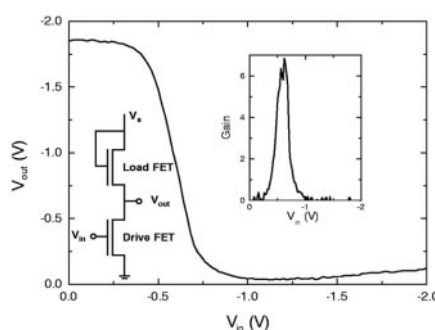
One of nanotechnology's rising stars is under investigation following claims that data in some of his papers have been falsified.

Jan Hendrik Schön, a researcher at Bell Labs in Murray Hill, New Jersey, faces an independent inquiry after scientists noticed striking similarities between different graphs in a number of his published papers.

Schön denies the allegations, but if they are upheld, many researchers are worried about the damage that could be done to basic research activities at Bell Labs, perhaps the world's best-known industrial laboratory. In addition, some fear that the allegations will taint nanotechnology's hot reputation.

The incessant drive towards miniaturization in electronics is a major theme of Schön's research. In particular, his work on using organic molecules deposited in thin films as molecular switches is viewed by many as a possible way to beat the size constraints imposed by silicon-based technologies.

The allegations against Schön first surfaced late last month after Lydia Sohn, who works on nanoscale electronics at Princeton University in New Jersey, received a tip-off from colleagues at Bell Labs. Sohn compared the graphs that make up Figure 4 in two separate papers published in *Science*¹ and *Nature*², and found the plots to be identical, right down to the random noise generated by the experiments (see above).



Double vision: similar graphs in *Science* (left) and *Nature* led some researchers to question the data.

Sohn notified the journals of the duplication, and they in turn requested that Schön provide clarification. He said that he had mistakenly sent the wrong figure to one of the journals and offered to write a correction.

Echoes from the past

That might have been the end of the matter had not one of Sohn's colleagues, Paul McEuen of Cornell University in New York state, made an unsettling discovery. "I was just looking at some of Schön's old papers and noticed a third similar figure," says McEuen. This graph was in a *Science* paper on a different type of microelectronic device³. The data were not identical to the first two, but portions of the graph matched perfectly.

A more extensive search by Sohn,

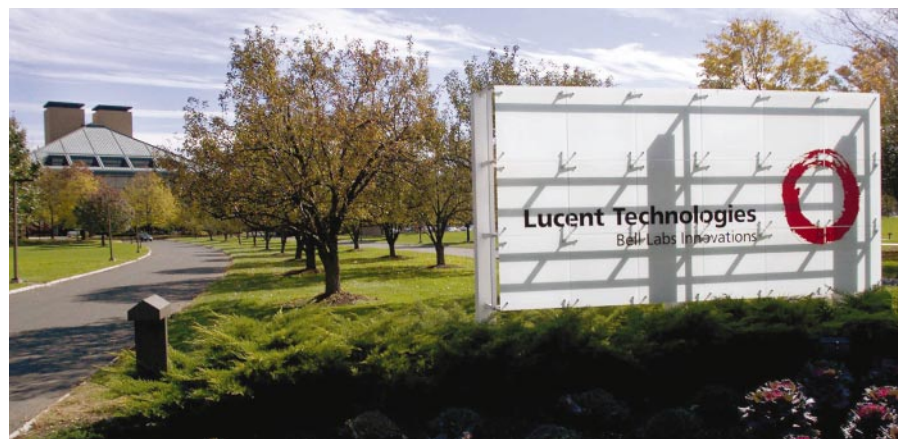
McEuen and others turned up a total of eight figures in six papers that appear to contain suspect data, the researchers say. In one case, two graphs look identical except for inverted scales, says Charles Lieber, a chemist at Harvard University. Lieber has examined all six of the papers under investigation, which include two others in *Science*^{4,5} and one in *Applied Physics Letters*⁶.

Senior officials at Bell Labs wrote to all the journals involved on 16 May saying that they had convened a special review committee, chaired by physicist Malcolm Beasley of Stanford University in California, to investigate the suspect figures. The investigation comes at a bad time for Bell Labs, as the financial difficulties of its parent company, telecommunications firm Lucent Technologies, have put pressure on the labs' research activities.

"We take concerns of scientific honour very seriously," says Cherry Murray, Bell Labs' senior vice-president of physical-sciences research, adding that Schön will be allowed to continue his work at Murray Hill until the review is completed. "We certainly don't want to rush to judgement," she says. Beasley hopes to complete the review by the end of the summer.

Nature will await the panel's findings before considering whether the authors should be asked to modify or retract the papers, says its editor, Philip Campbell.

Schön stands by his results. "I am confident in the measurements that I have taken and the experimental results, and I've tried to report them as best as I could," he says. Although he declines to comment on specific



Bell Labs at Murray Hill: launching an independent investigation into published nanotechnology results.



Rising star: Jan Hendrik Schön's impressive results have secured him several awards.

▶ allegations, Schön says that he is cooperating fully with the review committee. "I'm collaborating with my colleagues to reproduce these results and show them to the committee," he says. "I am trying to focus on the science."

At 31 years old, Schön is seen as one of the most able young physicists in nanotechnology. In the six years since he got his PhD at the University of Konstanz in Germany, he has produced over 100 papers and claimed several patents, as well as winning awards for his work in both the United States and Germany.

The speed and scope of his findings have aroused admiration among researchers — but some of his results have proved hard to reproduce. Robert Dynes at the University of California, San Diego, for example, has tried to replicate some of Schön's results for molecular switches that are turned on and off when an electric field is applied.

"I was fascinated by the results and frustrated that I couldn't reproduce them, and I didn't totally understand why I couldn't," recalls Dynes. The problem, he says, was that the applied electric field kept destroying critical components of the experiment.

Dynes is not alone. Groups at the French Atomic Energy Commission, Harvard University, Princeton and elsewhere say that they have so far been unable to reproduce some of Schön's results. ■

1. Schön, J. H., Meng, H. & Bao, Z. *Science* **294**, 2138–2140 (2001).
2. Schön, J. H., Meng, H. & Bao, Z. *Nature* **413**, 713–716 (2001).
3. Schön, J. H., Berg, S., Kloc, Ch. & Batlogg, B. *Science* **287**, 1022–1023 (2000).
4. Schön, J. H., Kloc, Ch., Haddon, R. C. & Batlogg, B. *Science* **288**, 656–658 (2000).
5. Schön, J. H., Dodabalapur, A., Kloc, Ch. & Batlogg, B. *Science* **290**, 963–965 (2000).
6. Schön, J. H., Kloc, Ch. & Batlogg, B. *Appl. Phys. Lett.* **77**, 3776–3778 (2000).

Public ponders biotech issues

Quirin Schiermeier, Munich

European citizens hold more finely differentiated and balanced views on genetically modified foods than scientists and politicians give them credit for, says a study carried out for the European Commission.

False interpretations of what the public wants are largely responsible for the difficulties faced by European policy-makers in managing agricultural biotechnology, the study claims.

The study, *Public Perceptions of Agricultural Biotechnologies in Europe*, is based on an analysis of discussions held between 1998 and 2000 by 55 focus groups in Germany, France, Britain, Italy and Spain, as well as interviews with activists, scientists and others who are more directly involved with the agricultural biotechnology industry. It was coordinated by Brian Wynne, a sociologist at Lancaster University, UK, who established his reputation by using such methods to assess the public's perception of nuclear power in the 1980s.

"Almost all popular opinions on the alleged misperceptions about the alleged view of the 'man and woman on the street' turned out to be simply myths," the study says. "Participants did not, overall, express entrenched opinions 'for' or 'against' genetically modified organisms."

The researchers found that public mistrust of regulatory bodies such as food-safety agencies was the underlying basis of suspicions of agricultural biotechnology.

Better information will not, in itself, restore the public's trust in these regulators,

the study concludes. Instead, it says that "more profound changes in institutional culture and practice" will be required. It suggests imposing heavy sanctions on companies or research institutions if any harm is caused by new technologies.

But Derek Burke, a retired molecular biologist and former chairman of the UK Advisory Committee on Novel Food and Processes, is sceptical about the study's findings, saying that he is uneasy about both its tone and its content. Burke argues that focus groups can easily be led towards a desired conclusion.

"This is an interesting contribution from a group of people with strong views," says Burke. "But their arguments reflect no more than the current media coverage." ■

♦ www.pabe.net



Food for thought: the public's perceived opinions on genetic modification are open to question.

US 'overspent' on collider project

Geoff Brumfiel, Washington

The US National Science Foundation (NSF) is under fire over its financial management of major research projects — including its contribution to the Large Hadron Collider, the particle collider being built at CERN, the European particle-physics laboratory in Geneva.

An audit by the research agency's inspector general, Christine Boesz, found that the NSF failed to track properly the full costs of the projects. For example, the agency told the National Science Board, its governing body, that its contribution to detectors at the Large Hadron Collider would cost \$81 million. But, the audit says, a further \$57 million will be needed for advanced computing and maintenance if US scientists are to glean any data from the detectors.

"The inspector general's report has confirmed my fears that there is little

oversight of NSF's large facilities," says Senator Kit Bond (Republican, Missouri), the senior Republican on the Senate subcommittee that funds the NSF. At a hearing on 15 May, Bond nonetheless expressed his support for a plan that would see the NSF's budget double in size within five years (see *Nature* 417, 209; 2002).

NSF officials claim that the inspector general's report is misleading. Robert Eisenstein, who heads the NSF's Directorate of Mathematical and Physical Sciences, says that the \$81 million will be enough to complete the parts of the detectors that are specified by the NSF's agreement with CERN. "We will deliver exactly what we said we will deliver," he says. The proposed \$57 million will cover maintenance and computing technology that had not even been invented when the agreement was signed, says Eisenstein. ■

Chemists synthesize a single naming system

David Adam, London

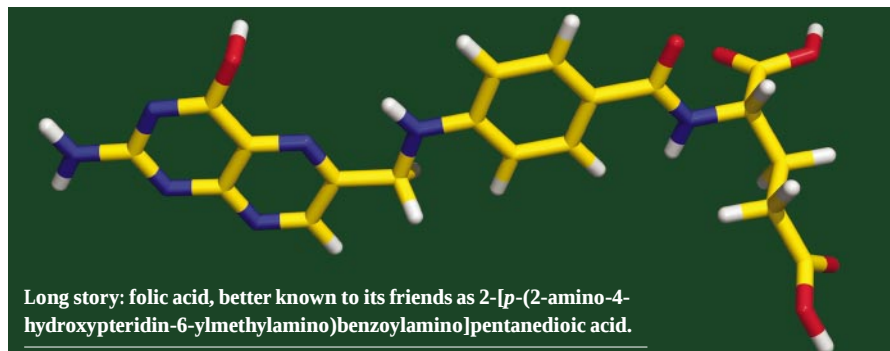
An international team of chemists is working on something that chemistry sorely lacks — a consistent and comprehensive way of labelling all chemical compounds.

The new technique will apply computer algorithms to molecular structures, generating a unique digital signature for any chemical compound. The new labels are not intended to replace common chemical names, but to allow easier linking to compounds in online chemical databases and journals.

"The hope is that all organizations that handle information on chemicals will be able to use a single format to say what the chemical is," says Alan McNaught, general manager of the production division at the Royal Society of Chemistry in Cambridge, who coordinates the project for the International Union of Pure and Applied Chemistry (IUPAC).

Right now there is no single international standard for identifying chemicals. The IUPAC and the American Chemical Society use different rules. Some drug companies, as well as different branches of chemistry, have their own chemical-naming systems. Even simple structures can cause confusion. For example, the formal name for acetic acid, the main ingredient in vinegar, is ethanoic acid.

IUPAC believes that its new system — which would be freely available to all — could unify the different approaches. Tentatively



Long story: folic acid, better known to its friends as 2-[p-(2-amino-4-hydroxypteridin-6-ylmethylamino)benzoylamino]pentanedioic acid.

known as IChI, for IUPAC chemical identifier, its development is being led by a team at the US National Institute of Standards and Technology (NIST) in Gaithersburg, Maryland.

A preliminary version of the software covering well-defined, covalently bonded organic molecules was released this year to let other chemists test the idea. It labels each atom in a compound in a way that does not depend on how the structure is drawn, and converts the label to a string of characters. The format has not been finalized, but at present ethane is 'C3C3, 2-1', for example, and acetone is 'C3C3OC, 4-1-2-3' (the labels are easily converted to structures using the algorithm). The process is reversible, so molecular structures can be generated from the identifiers.

The next step is to extend the system to include more complex organic compounds,

such as polymers, and ultimately to tackle inorganic compounds. By adding it to software packages commonly used to draw chemical structures, the NIST team hopes that IChI will enter into widespread use.

In effect, the IChI number will provide each chemical molecule with a digital object identifier (DOI) — a concept increasingly being applied to everything from scientific papers to individual genes. Jonathan Goodman, a chemist at the University of Cambridge, says chemistry suits this approach well. "Molecules are a wonderful unit of information to treat in this way," he says. "They are complex enough to have lots of interesting features and difficulties but simple enough to represent quite a small subset." ■

♦ www.iupac.org/projects/2000/2000-025-1-800.html

NASA's critics demand a few more steps into space

Tony Reichhardt, Washington

Critics of Sean O'Keefe, the new NASA administrator, are complaining that his nuts-and-bolts management style is leaving the space agency bereft of what ex-president George Bush once called "the vision thing".

The agency's human spaceflight programme is "adrift, with no clear vision or commitment to any goals after the completion of the International Space Station", Congressman Nick Lampson (Democrat, Texas) told the House of Representatives on 15 May.

Lampson was introducing the Space Exploration Act of 2002, which would require NASA to complete a series of specific tasks, starting with building a reusable space vehicle within eight years that is capable of carrying astronauts beyond Earth orbit to locations where large space telescopes could be constructed. Under the act, NASA would also be charged with building, within 10 years, a vehicle capable of reaching near-Earth asteroids. A lunar research facility would follow within 15 years, and a spacecraft for reaching Mars within 20.



The bill takes an approach that is becoming fashionable among long-term space planners: it proposes a series of incremental steps, rather than the grand sweep of a single mission to Mars. It calls for a competitive programme, run by a new Office of Exploration, modelled on NASA's Discovery series of planetary-exploration missions, with an initial budget of \$50

million next year and \$200 million in 2004.

Lampson's proposal contrasts sharply with the approach being taken by the White House and O'Keefe, who has made it clear that he envisages a period of incremental technology development for NASA, rather than identifying any grandiose new goals.

Although Lampson's bill has little chance of passing Congress, NASA's recent 'go-slow' strategy has been criticized by other politicians, including Tom DeLay (Republican, Texas), the majority whip in the House and one of Washington's most powerful figures. During a congressional hearing last month, DeLay branded O'Keefe's plans for human spaceflight "timid and anaemic".

Most of the bill's sponsors are from districts near NASA's Johnson Space Center in Texas, where the US astronaut programme is based. But its arrival, observers say, reflects a growing weight of opinion that, with the space station downsized and no other large projects on the horizon, NASA's 40-year-old human-spaceflight programme is in danger of withering away. ■

Think-tank calls for an end to DNA deception

David Adam, London

Deceitfully obtaining and testing someone else's DNA would become a criminal offence under recommendations delivered to the British government by a high-powered commission.

The Human Genetics Commission says it is worried that no law currently prevents DNA samples of the rich and famous being obtained, analysed and published — or stops unscrupulous family members from taking material from children and secretly carrying out paternity tests.

The recommendation is one of several made in a report, *Inside Information*, that the commission was due to release on 21 May. The government has welcomed the report and says that it will consider the issues raised during the government's discussions on genetics, the results of which it expects to publish later this year.

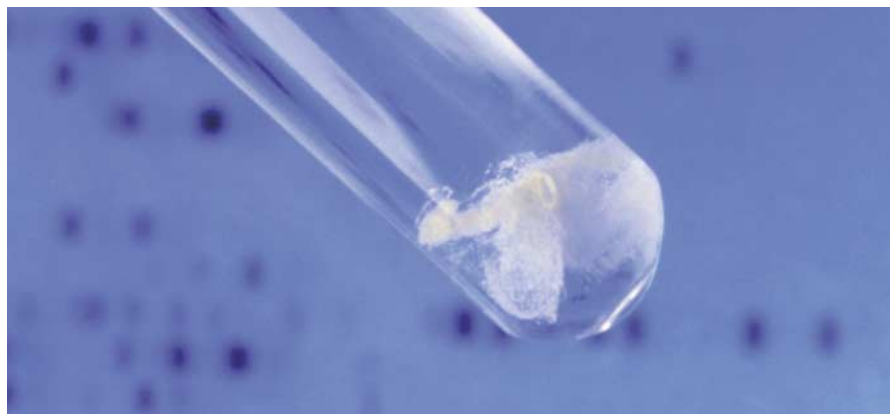
"There must be systems in place that promote trust about the way that clinicians, researchers and ultimately the state handle genetic information," it says. "People need to feel confident that their information is safe from being used in the wrong way." To address these concerns, it calls for wider debate over ways to control the use of genetic information in police forensic science, clinical research projects and setting personal insurance premiums.

Civil-liberties groups in Britain are particularly worried by a recent case in which an HIV-positive prisoner volunteered a blood sample for a genetic study, only to see the sample taken by police and used successfully to prosecute the man for deliberately transmitting HIV to an ex-girlfriend.

Some researchers are worried that similar cases will ultimately deter prisoners and others from consenting to participate in medical trials or in projects such as the Medical Research Council's Biobank UK, which hopes to collect DNA samples from half a million people (see *Nature* 417, 9; 2002). The commission says that a law may be needed to bar police and other officials from access to genetic research databases.

"We've tried to identify the main principles which should govern the way our society deals with personal genetic information," says Alexander McCall Smith, a professor of law at the University of Edinburgh who specializes in medical and criminal law and is vice-chairman of the Human Genetics Commission. Many of the techniques and issues considered are only at very early stages, but it is important to establish a regulatory framework of legislation and guidelines before they develop further, he says.

The commission says that public concern may arise as growing knowledge of genetics



Concerns are growing about covert use of DNA samples (above) for forensic and insurance purposes.

allows aspects of a suspect to be built up from samples found at a crime scene — or even to predict elements of the suspect's behaviour or medical history. McCall Smith points out, for example, that the forensic-science service can already tell the police the chances of a suspect having ginger hair based on the analysis of just one gene, that encoding melanocortin receptor 1.

Paul Debenham, director of life sciences

at the Laboratory of the Government Chemist, based in London, which provides genetic-analysis services for the police and government, says he would welcome guidelines covering the forensic uses of DNA. Experts say, however, that legislation to prevent authorities obtaining medical information from DNA would be difficult to craft in such a way as to allow the collection of data related to suspects' appearances. ■

Biotech firm's accounts scrutinized

Erika Check, Washington

A Massachusetts biotechnology company that usually likes to keep a high profile received some unwanted publicity this week when government auditors raised questions about its financial viability.

Auditors from the health department were asked to trawl through the accounts of Advanced Cell Technology (ACT) of Worcester by Congressman Joe Pitts (Republican, Pennsylvania), who is strongly opposed to human cloning. He says that he wanted to find out whether the company had illegally used federal money

from the National Institutes of Health (NIH) when it created cloned human embryos in November last year (see *Nature* 414, 477; 2001).

The auditors say that they found "no evidence that NIH's funds supported ACT's embryo-cloning activities". But they did find that ACT's accounting system was inadequate, and that

the company had made "unallowable" equipment purchases with the grant money.

They also found what the published audit calls "concerns regarding ACT's financial viability". And the audit raises what it terms "concerns regarding the continuity of NIH-funded research" under two NIH grants totalling \$1.8 million, since the grants' principal investigators left the company this spring after one year of service with the company.

ACT says it had done nothing wrong, but has agreed to return the funds in question to the government, in line with the auditor's recommendations.

Michael West, president of ACT, accuses Pitts of making misleading statements about his company on the basis of the audit. "After weeks of looking at every scrap of paper in the company, they found no evidence that we misappropriated grant funds for human embryo research," he says. "So they tried to make it look like we were misappropriating grant money in other ways — and it's not true."

ACT is in no worse shape financially than many other similar companies, West says. "We're clearly a small, struggling biotechnology company," he says, "but we're doing better than most." ■



Joe Pitts called for ACT to be audited.

Canadian collaboration aims to protect tribes

Rex Dalton, San Diego

Ethnobotanist Kelly Bannister feels strongly about protecting the rights of her research subjects in a native tribe in Western Canada — so strongly that she has previously put them ahead of her own rights to publish her work.

Now she is extending her philosophy into a programme called Community-University Connections. Already engaged in its first project, at the Clayoquot Biosphere Trust in British Columbia, the programme aims to help improve research collaborations between indigenous tribes and researchers.

The seeds of the idea were sown when Bannister was doing her doctoral thesis at the University of British Columbia in Vancouver. Guided by the elders of the Skeetchestn tribe, Bannister was studying plants with historic medicinal uses. In 1997, her academic adviser wanted to set up a collaboration with a drug company to study potential antimicrobial compounds identified by her research. But the tribe needed time to record and publish the knowledge of their elders, so Bannister denied her adviser access to the plant-assay data she had collected.

And in 2000 when she completed her doctorate, under a different supervisor, Bannister agreed to seal her thesis for five years to give the tribe time to prepare claims to any compounds that might be commercialized from the 70 plants she had analysed.

"I felt the data were going to be taken over

and used for a purpose not in the best interest of the indigenous community," says Bannister. "I decided to take a stand; change requires people to stand up."

Now a postdoctoral fellow at the University of Victoria in British Columbia, Bannister is developing the Community-University Connections collaborative model for other researchers. She notes that the best interests of indigenous peoples often collide with the eagerness of most university researchers to publish information quickly. "Sometimes publications are not the best for the indigenous community," she says.

Ron Ignace, chief of the Skeetchestn tribe near Kamloops, British Columbia, claims that his people have been abused in the past by researchers who published their tribal history without permission. "Fortunately, this time we had someone with high moral fibre who worked with us," he says.

The tribe now plans to develop an ethnobotany research laboratory at the community's university campus. But with the lab still being planned and the data still under seal, work has yet to begin on Bannister's data.

Bannister and her colleagues agree that her approach can slow research projects, but they say that the results are worth the wait. Rod Dobell, a Victoria professor of public policy involved in the Community-University project, says that such negotiations "at the front end can prevent paralyzing roadblocks later".

Jo Render, a programme officer at First

Nations Development Institute, a Virginia group that fosters indigenous rights, says that Bannister's developing model shows great promise. "If people are willing to defend their thesis on the scientific results, why shouldn't they be willing to defend the process they use?" Render asks.

♦ <http://web.uvic.ca/~scishops>



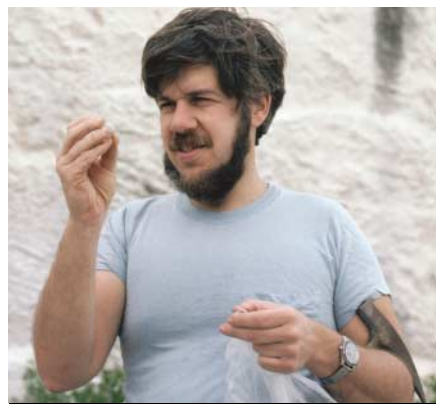
An ethical route: Kelly Bannister wants to see indigenous peoples benefit from research.

D. MUENCH/CORBIS

Harvard mourns loss of popular evolutionary thinker

Erika Check, Washington

Stephen Jay Gould, perhaps the best-known evolutionary theorist since Charles Darwin, died of lung cancer on 20 May. He was 60 years old and had been fighting against cancer since being diagnosed with mesothelioma two decades ago.



Stephen Jay Gould took evolution to the masses.

Gould, who spent almost his entire career at Harvard University, was famous for his outspoken views and wide-ranging intellect. He came to scientific attention in the early 1970s by formulating the controversial theory of punctuated equilibria — which argues that evolution tends to occur in bursts, interspersed with long periods of relative stasis — with Niles Eldredge of the American Museum of Natural History in New York.

Gould's own research concentrated on the palaeontology and evolution of West Indian land snails, but it was through his elegant writing about evolution and other scientific mysteries that he really made his mark. He won acclaim for popular books such as *The Panda's Thumb* (1980), which won the American Book Award for Science, *The Mismeasure of Man* (1981), which won the National Book Critics Circle Award, and *Wonderful Life* (1989).

"Steve did not try to make it simple, he tried and succeeded in explaining the

complications," says evolutionary geneticist Richard Lewontin, one of many Harvard colleagues to pay tribute. "He made readers appreciate how messy and variable life is."

Most recently, Gould published a massive intellectual tome, *The Structure of Evolutionary Theory*, a work he began more than 20 years ago (see *Nature* 416, 787; 2002). Gould also campaigned to ensure that evolution would be taught in US schools.

Born in New York City, Gould earned a bachelor's degree in zoology from Antioch College in 1963 and a PhD from Columbia University in 1967. He then took a post at Harvard, eventually becoming the Alexander Agassiz Professor of Zoology in 1982.

Gould was also known for his nonscientific pursuits. He wrote enthusiastically about baseball and sang for many years with a Boston-area choir. He is survived by his second wife and two children from a previous marriage.

♦ www.hno.harvard.edu/gazette/2002/05.16/99-gould.html

W. MCNAMEE/CORBIS

Chile boosts science by joining European research network

Washington Chile's president Ricardo Lagos signed a free-trade agreement with the European Union (EU) on 17 May, making Chile the first Latin American country to join the EU's Sixth Framework Programme.

Chile's entry into the 16.2-billion-euro (US\$14-billion) programme will boost its small but growing science base. Eric Goles, president of Chile's National Commission on Scientific and Technological Research, says that a top priority will be to fortify the country's Centres of Excellence, which are dedicated to molecular biology, mathematics, biotechnology, materials science, oceanography and astronomy.

The five-year agreement, due to start in 2003, may help Chile towards its stated goal of doubling expenditure on research and development from 0.6% to 1.2% of its gross domestic product.

German animals win legal right to state protection

Munich German researchers who use animals in their work face a new round of legal battles after the country's parliament voted to give

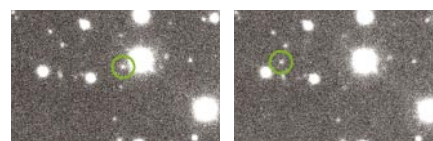
animals a constitutional right to be protected by the state.

The new amendment, passed last week by 543 votes to 19, states that animals must be treated as fellow creatures and protected from avoidable pain. German researchers had lobbied against the move, fearing that it would interfere with the constitutionally guaranteed freedom to research. Animal-welfare groups say they will consider using the amendment in legal attempts to reduce the number of animal experiments carried out in Germany (see *Nature* 416, 355; 2002).

The change follows a controversial legal judgement on halal slaughtering. In January, Germany's highest court ruled that Muslims should be allowed to kill animals by the halal method, in which the animal is not stunned beforehand. The decision raised the level of debate about animal rights and prompted the Christian Democrats, who had previously opposed changing the constitution, to back the amendment.

New moons for Jupiter leave Saturn eclipsed

Hawaii Jupiter regained its position as the planet with the most known moons last week. Astronomers from the University of Hawaii announced the discovery of 11 new satellites, bringing Jupiter's total to 39.



Good spot: a new moon (circled) of Jupiter was seen because it moved relative to the background.

Saturn has 30 known moons, a dozen of which were found just last year.

Scott Sheppard, David Jewitt and colleagues used the modest 3.6-metre Canada–France–Hawaii Telescope on Mauna Kea to survey the region around Jupiter last December. Computers were then used to scan the images for objects that were moving against the background of stars and galaxies. The researchers found 11 small objects 2–4 kilometres in size, and were able to pin down the objects' irregularly shaped orbits.

The moons were probably pulled into orbit around Jupiter early in its history, when they were slowed down by the gravitational attraction of Jupiter or its atmosphere.

Cancer chemical found in fried and baked food

London Two studies in the past month have revealed the presence of a potential cancer-

causing chemical in food that has been fried or baked.

A study carried out for the Swedish National Food Administration, published on 24 April, found that foods such as potato chips and bread contain a chemical called acrylamide. The effect of acrylamide on human health is unclear, but it has been shown to induce cancer in animals, and is classified as “probably carcinogenic to humans” by the World Health Organization (WHO). The Swedish findings were supported by a second report, released by Britain’s Food Standards Agency on 17 May.

Further research into the effects of acrylamide will be discussed at a WHO meeting at the end of June.

Transgenic crops may seed growth in costs

Paris Introducing large-scale farming of genetically modified (GM) crops could drive up the prices of non-transgenic varieties, a new report from the European Commission has warned.

The price increases will come from changes to farming practices that will be needed to ensure that the level of GM ingredients is kept below certain thresholds in food that is labelled as GM-free. Buffer

zones will have to be created around transgenic crops to guard against cross-pollination, for example, and farmers will have to monitor non-GM crops to check for contamination. The report says that these measures could lead to an increase of between 1% and 10% in the price of potatoes and maize, and up to 41% for organic oilseed rape seeds grown from seeds saved from previous years’ crops.

Scientific opinion and computer simulations of different farming practices were used to model the coexistence of GM, conventional and organic crops in European agriculture. But data are limited because transgenic crops are not currently grown on a large scale in Europe.

Whale of a problem for vets as orphaned orca gets sick

Seattle The blubber is flying over whether to rescue an orphaned killer whale, or orca, near Seattle. The female yearling arrived in the Puget Sound in January after being separated from the rest of her pod. She has since developed a severe skin condition, a parasite infestation and breath that smells like paint-thinner.

The Vancouver Aquarium offered in March to capture and rehabilitate the killer whale so that she could rejoin her pod when



Kill or cure? This killer whale's bad breath could be an indication of a metabolic illness.

they return in July, a plan that is endorsed by marine-mammal experts. But the US National Marine Fisheries Service says that it opposes capture because of the stress involved.

Vets at the fisheries service believe that the whale's bad breath is a sign of a metabolic illness, and say that she should not be moved until it is clear what is wrong. But blood tests for such a condition were negative last week, and the service's veterinary panel will reconvene this week.

Rescue proponents say that they cannot understand the foot-dragging. The aquarium's president, John Nightingale, says the delays have dashed any hope of treating the whale in time for a rendezvous this summer. “We spent about \$20,000 on this and all we got in return was a kick in the shins,” he says.

B. WOOD/PROJECT SEAWOLF



Common ground: relationships between different species are clearly written in their genes, but not in their chromosomes.

All genomes great and small

Chromosome size and number can vary widely between closely related organisms. This poses a challenge for the evolutionary geneticists who are trying to make sense of genome structure, says Jonathan Knight.

We are somewhat like bananas, at least as far as our genes are concerned. About a third of human genes are clearly related to those found in plants. Fruitflies, sitting closer to us on the tree of life, share about two-thirds of our genes. And mice, being fellow mammals, are at least 90% genetically similar to us.

But look at the size and structure of the chromosomes on which those genes sit, and the neat correlation between evolutionary relatedness and genomic similarity breaks down. Sheep have 27 pairs of chromosomes; the Indian muntjac deer has just 3. We have some 3 billion DNA base pairs; one species of amoeba has more than 600 billion.

It seems entirely random. But as the genomes of ever more organisms are sequenced in their entirety, trying to make sense of natural variability in genome structure has become a burgeoning field. "We're finally able to ask some key questions," says Daniel Hartl, an evolutionary geneticist at Harvard University.

Although it is still early days, comparisons of sequences from different species suggest that events such as bursts of activity of 'jumping genes', genetic duplications and chromosome fusions play a key role in evolution. Far from being a mass of junk DNA that holds a small but precious cargo of genes, geneticists are starting to see chromosomes as highly dynamic stages on which important evolutionary processes are played out.

Jump to it

The mobile elements known as transposons are among the most powerful forces that shape chromosome evolution. These 'jumping genes' carry instructions for their own excision, duplication and insertion into the genome. It seems that, at certain periods in evolutionary history, transposon activity made chromosomes expand like accordions. As a result, most chromosomes are now filled with the silent remnants of transposons.

Evolutionary geneticists can work out how long ago a transposon became immobi-

lized by looking at the accumulation of mutations in its characteristic flanking sequences. Such studies have suggested that a flurry of transposon activity doubled the size of the maize genome from 1.2 billion to 2.4 billion bases in the past 3 million years¹. In human evolution, transposons called long interspersed elements (LINEs) have expanded to about 100,000 copies in several bursts over the past 100 million years, the most recent event having occurred 25 million years ago². LINEs now account for 15% of human DNA.

Why should mobile elements disperse themselves in bursts? One intriguing idea, says Hartl, is that most of the time cells repress transposon activity — a sensible strategy, given that a gene can be disabled if a transposon jumps into its sequence. But the costs and benefits may shift during periods of greater evolutionary stress. Increased rates of transposition might then be selected for, Hartl argues, because they may help organisms to adapt in tough times by increasing genetic variability.

A. SYRED/SPL; M. S. YAMASHITA/CORBIS; G. DOUWMA/NPL; CORBIS; D. DALE/SPL

	Genome size (base pairs)	Chromosome number (<i>n</i>)
<i>Amoeba dubia</i>	670,000,000,000	Several hundred
Trumpet lily (<i>Lilium longiflorum</i>)	90,000,000,000	12
Mouse (<i>Mus musculus</i>)	3,454,200,000	20
Human (<i>Homo sapiens</i>)	3,200,000,000	23
Carp (<i>Cyprinus carpio</i>)	1,700,000,000	49
Chicken (<i>Gallus gallus</i>)	1,200,000,000	39
Housefly (<i>Musca domestica</i>)	900,000,000	6
Tomato (<i>Lycopersicon esculentum</i>)	655,000,000	12

What's in a number? Genome size and chromosome number seem unrelated to complexity. The figures above are for haploid genomes — most cells are diploid (2*n*), carrying two copies of each chromosome.

Transposons aren't the only type of DNA that can be duplicated. If there is one thing DNA is good at, it is being copied. And as a result, duplications ranging from hundreds of bases to the cell's entire complement of chromosomes have figured heavily in the evolution of modern genomes.

The simplest duplications produce adjacent repeated sequences that are all oriented in the same way. The lengths of these 'tandem repeats' can vary greatly, and often entire genes are duplicated. The extra copy can then accumulate mutations; often these will render it useless, but a new and useful function can also emerge. Gene duplication is now widely considered to be the most likely origin of gene clusters, in which natural selection has shaped copies of one original gene to take on different functions. We owe our wide-ranging sense of smell, for instance, to the duplication and diversification of olfactory-receptor genes³.

Duplicated DNA need not end up close to its template. The human genome contains duplicated chunks of DNA hundreds of kilobases long at opposite ends of a chromosome or on different chromosomes altogether⁴. Evan Eichler, a genome researcher at Case Western Reserve University in Cleveland, Ohio, estimates that at least 5% of the human genome arose through this sort of duplication⁵. Indeed, the true figure may be much higher, as duplications in an ancestor that lived more than 40 million years ago would by now be undetectable because of the subsequent divergence of the copied regions.

Eichler and his colleagues have found that more than a third of duplicated regions that end up on another chromosome are found near the centromere, the anchor point for the

transcribed into RNA. But some centromeres in the human genome have an intermediate region between the satellite DNA and the adjacent, less closely packed 'euchromatin' that houses active genes. Eichler's group has found that duplicated segments may pop up as islands in a sea of α -satellite DNA in a region known as the pericentromere⁶.

Some pericentromeres are part vacuum cleaner, part blender. They suck in duplications, chop them up and recombine them. As new insertions arrive, they drop in without respect for other insertions, sometimes splitting them in two. Once in a pericentromere, sections of duplicated DNA are clipped out, inverted and stuck in elsewhere. The result is a tossed salad of duplications — which Eichler suspects may serve as an incubator of new genes.

Garbage or gold?

Many of the duplications have brought with them the regulatory sequences that allow them to be transcribed, and Eichler's team has detected RNA messages from pericentromeric regions⁷. Although most of the resulting transcripts are garbage, there is the possibility that one could occasionally code for a new and useful protein. So far, no example of a new gene being born this way has been found. But Eichler is examining the regions around a 'dead' centromere on human chromosome 2. Here there are higher rates of transcription than around the chromosome's active centromere, and Eichler hopes to find new and interesting genes.

Other researchers are studying analogous regions of repetitive DNA that sit adjacent to the telomeres that cap chromosomes' ends. Like pericentromeres, these 'subtelomeres'

seem to be zones of active genetic recombination. Researchers led by Barbara Trask of the Fred Hutchinson Cancer Research Center in Seattle have shown that sections of subtelomere some 50–100 kilobases long have moved from one chromosome to another in our species'

recent evolutionary history⁸.

Trask has also found several members of the family of olfactory-receptor genes hiding in human subtelomeres, at least one of which seems to be functional, being expressed in the olfactory epithelium⁹. Together, her discoveries suggest that, like pericentromeres, subtelomeres may be important regions for the mixing and matching of duplicated DNA to form new genes.

Intriguingly, in earlier studies carried out at the Lawrence Livermore National Laboratory in California, Trask found that most of the variability in overall genome size across human populations is due to variation in the sizes of pericentromeres and subtelomeres¹⁰. In the case of chromosome 21, this variation makes the chromosome 45% larger in some people than in others. The expansion of subtelomeres could be functionally and evolutionarily significant, Trask suggests.

Doubling up

Such duplication and recombination events cannot explain the wide variety in chromosome number seen in nature. But individual chromosomes, and even entire sets, can also be duplicated. More than three decades ago, Susumu Ohno, a geneticist at the City of Hope Cancer Center in Los Angeles, proposed that genome duplication could account for the relatively rapid evolution of complexity in vertebrates¹¹.

Most organisms are diploid — they carry two copies of each chromosome, one from their father, one from their mother. But an error in chromosome segregation can easily result in an individual being tetraploid, carrying twice the usual complement of chromosomes. Ohno argued that this was a key factor in vertebrate evolutionary history. Over time, mutations would accumulate in the duplicated chromosomes until they were so divergent that they were clearly distinct. The organism would then seem to be diploid once more — only now it would have twice the number of chromosomes.

Examples of genome duplication are easy enough to find, particularly in flowering plants, where only about half of all species are diploid. Wheat, for example, has six copies of each chromosome. And when plant biologists examined the genome sequence of the diploid

thale cress (*Arabidopsis thaliana*), they concluded from extensive evidence of duplication that it must have evolved from a tetraploid ancestor that first arose around 112 million years ago¹².

Similar polyploidy almost certainly figured in the evolution of the yeast *Saccha-*



Daniel Hartl thinks deletion determines genome size.

protein filaments that yank freshly divided chromosomes apart during cell division. Centromeres consist mainly of tandem repeats, 171 bases long, known as α -satellite DNA. This DNA tends to remain tightly packed in a structure called heterochromatin, where it is not



The yeast genome was shaped by duplication, says Kenneth Wolfe.



Evan Eichler is looking for new genes in a duplication salad.



The Indian (left) and Chinese (right) muntjacs have ended up with very different chromosome counts.

Saccharomyces cerevisiae, adds Kenneth Wolfe, a geneticist at Trinity College Dublin. In 1997, he reported that the sequences that flank the centromeres of *S. cerevisiae*'s 16 chromosomes are actually grouped into 8 pairs¹³. "What we are looking at are the evolutionary products of some kind of mistake in chromosome pairing that happened millions of years ago," says Wolfe. Although others have argued that Wolfe's observations could be the result of multiple duplications, Wolfe says that his unpublished analysis of other regions of the yeast genome suggests that it was a single, dramatic event.

A theory with backbone

But what of Ohno's theory that genome duplications were critical to the evolution of vertebrates? The idea gained support in the late 1980s with the discovery in mice and humans of four nearly identical gene clusters that control the development of the body plan, known as the homeobox, or Hox, genes. Fruitflies — in which Hox genes were first discovered — and other invertebrates have only a single cluster, but the order of the genes within it is the same as in the mammalian clusters. So human Hox genes seem to be the product of two rounds of duplication of an ancestral Hox cluster¹⁴.

When geneticists discovered several other gene clusters that appear just once in invertebrate genomes but in quadruplicate in vertebrates, a consensus emerged that the entire genome must have been duplicated twice in some early ancestor within the vertebrate lineage¹⁵. But now that the human genome has been sequenced in draft form, the debate has opened up once more.

Austin Hughes, an evolutionary biologist at the University of South Carolina in Columbia, has found that fewer than 5% of human genes that have invertebrate homologues appear in quadruplicate. Furthermore, of 134 regions that do have four copies, only 30% are organized into two clusters of two, as would be expected if two successive genome duplications had occurred early in vertebrate

evolutionary history¹⁶. "To me that closes the book on it," says Hughes.

Not so for Wolfe. Mammalian and bird embryos cannot survive if they are tetraploid, so any genome duplications in the vertebrate lineage must have taken place more than 200 million years ago, before either of these groups had evolved. Such an ancient event would be extremely hard to spot in today's genomes, Wolfe argues. For one thing, many duplicate genes — possibly most of them — may have been lost. In yeast, where evidence of polyploidy is stronger, only 16% of genes seem to have a surviving homologous partner¹⁵. So for Wolfe, whole-genome duplication may still have played an important role in vertebrate evolution.

Species can also experience sudden reductions in chromosome number, as is clear from studies of muntjacs, a diminutive genus of deer. The Chinese muntjac (*Muntiacus reevesi*) has 23 pairs of chromosomes, but the Indian muntjac (*M. muntjac*) has only 3 pairs¹⁷. Other members of the genus have intermediate numbers of chromosomes. Yet the banding patterns of the chromosomes in all muntjac species suggest that they contain roughly the same genes.

Shrinking genomes

Wen Wang and Hong Lan of the Chinese Academy of Sciences in Kunming, Yunnan Province, have produced an evolutionary tree of the seven muntjac species on the basis of the DNA found in their energy-generating mitochondria. Their results suggest that the ancestral muntjac species had a high chromosome number, which decreased in some lineages as a result of end-to-end fusions between different chromosomes¹⁸. And Fengtang Yang and Malcolm Ferguson-Smith at the University of Cambridge, UK, have used the colourful technique of chromosome painting to show that the Indian muntjac chromosome 3 is an assemblage of seven chromosomes from its Chinese cousin¹⁹.

Chromosomes can decrease in size as well

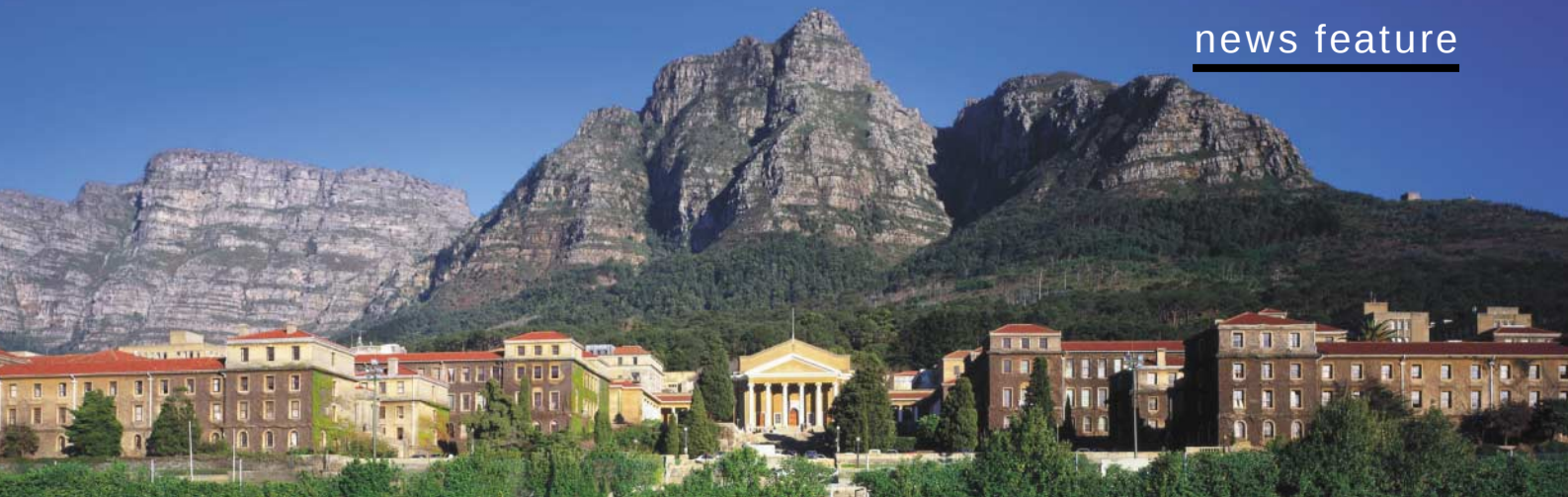
as in number — indeed, if there wasn't some shrinking, duplication events would cause genomes to grow inexorably over evolutionary time. In fact, Hartl suggests that spontaneous deletion may be the principal process that determines genome size. His team estimated the rate of deletions in several organisms by looking for evidence of excisions within dead transposons, the ages of which can be estimated by the accumulation of mutations in their flanking sequences. Hartl found that the fruitfly *Drosophila melanogaster* is losing DNA 60 times faster than mammals, which fits with its much smaller genome size. Hawaiian crickets (*Laupala* sp.), which have 11 times more DNA than fruitflies, lose DNA 40 times more slowly. And the grasshopper *Podisma pedestris*, whose genome is 100 times bigger than that of *Drosophila*, loses DNA more slowly still²⁰.

Evolutionary geneticists are now pondering the significance of the processes that have shaped modern genomes. The more they learn, the less those vast stretches of non-coding DNA that litter most genomes seem like just junk. Pericentromeres and subtelomeres may be incubators for new genes. Duplications may help to generate genetic variation. And natural selection may continually force genomes to become leaner and fitter.

Undoubtedly these processes have also left behind heaps of rubbish. But just as the surface of the Moon holds a record of its impact history in its craters, those genetic wastelands can serve as records of the dynamic history of individual genomes. Now equipped with the best maps yet, the explorers are setting out.

Jonathan Knight writes for *Nature* from San Francisco.

- San Miguel, P., Gaut, B. S., Tikhonov, A., Nakajima, Y. & Bennetzen, J. L. *Nature Genet.* **20**, 43–45 (1998).
- Sheen, F. M. et al. *Genome Res.* **10**, 1496–1508 (2000).
- Glusman, G., Yanai, I., Rubin, I. & Lancet, D. *Genome Res.* **11**, 685–702 (2001).
- The International Human Genome Sequencing Consortium. *Nature* **409**, 860–921 (2001).
- Samonte, R. V. & Eichler, E. E. *Nature Rev. Genet.* **3**, 65–72 (2002).
- Bailey, J. A., Yavor, A. M., Massa, H. F., Trask, B. J. & Eichler, E. E. *Genome Res.* **11**, 1005–1017 (2001).
- Bailey, J. A. et al. *Am. J. Hum. Genet.* **70**, 83–100 (2002).
- Mefford, H. C. & Trask, B. J. *Nature Rev. Genet.* **3**, 91–102 (2002).
- Linardopoulou, E. et al. *Hum. Mol. Genet.* **10**, 2373–2383 (2001).
- Trask, B. J., van den Engh, G., Mayall, B. & Gray, J. W. *Am. J. Hum. Genet.* **45**, 739–752 (1989).
- Ohno, S. *Evolution by Gene Duplication* (Springer, Berlin, 1970).
- The Arabidopsis Genome Initiative. *Nature* **408**, 796–815 (2000).
- Wolfe, K. H. & Shields, D. C. *Nature* **387**, 708–713 (1997).
- Schughart, K., Kappen, C. & Ruddle, F. H. *Br. J. Cancer Suppl.* **9**, 9–13 (1988).
- Wolfe, K. H. *Nature Rev. Genet.* **2**, 333–341 (2001).
- Friedman, R. & Hughes, A. L. *Genome Res.* **11**, 1842–1847 (2001).
- Liming, S., Yingying, Y. & Xingsheng, D. *Cytogenet. Cell Genet.* **26**, 22–27 (1980).
- Wang, W. & Lan, H. *Mol. Biol. Evol.* **17**, 1326–1333 (2000).
- Yang, F., O'Brien, P. C. M., Wienberg, J. & Ferguson-Smith, M. A. *Chromosome Res.* **5**, 109–117 (1997).
- Hartl, D. L. *Nature Rev. Genet.* **1**, 145–149 (2000).



UCT The rainbow academic nation

South Africa's higher-education system was designed by the architects of apartheid. So why are today's academics resisting attempts to reform it? Michael Cherry investigates.

Amid the grim news of AIDS and crime that flows out of South Africa, one success story has gone largely unreported. During the apartheid era, higher education, like most elements of South African society, was divided along racial lines. Now, less than 10 years after the country's first open elections, white students account for less than half of the enrolments at institutions that were once exclusively white.

But successful integration has brought fresh problems. As black students — a term that in South Africa encompasses black Africans, Indians and those of mixed race — abandon the universities that were created for them, the government has been forced to consider merger plans. And these are proving far from popular with both historically black and historically white institutions. Places such as the historically black University of Fort Hare, which counts Nelson Mandela among its alumni, complain that their unique identity will be lost. As the South African cabinet ponders the future of its higher-education system, it is far from clear whether the successes of the 1990s can be maintained.

Current figures for enrolment in South African higher education paint an encouraging picture. Black Africans, who make up about three-quarters of the country's population, account for 60% of student numbers. Representation of mixed-race (5%) and Indian (7%) students also mirrors the country's racial diversity. And an increasing number of black Africans are going to historically white schools — 43% in 2000, compared with 13% in 1993, the year before the first free elections.

This change has meant that both institu-

tions and students have had to adapt. "To change the profile and maintain standards is a tough call," says Nicola Illing, head of molecular and cell biology at the University of Cape Town. "Disadvantaged students are usually inadequately prepared." In addition, some of the first black students to join predominantly white universities felt somewhat uncomfortable. But despite such difficulties, academics are encouraged by the progress that has been made. Half of all the postgraduates in Illing's department, for instance, are now black.

Faded glory

But it is the historically black institutions that face the biggest problems. Some of these universities have long histories, but their reputations suffered during apartheid, when they were designated as centres for specific ethnic groups. Unable to attract the best staff, many suffered from poor management and few performed any research of note.

As the racial barriers to education were slowly removed during the 1970s and 1980s, the brightest students opted — if they could afford to — for historically white institutions. The trend accelerated with the dismantling of apartheid, and enrolment in historically black universities and technikons — which focus on vocational training and perform little research — fell steadily during the 1990s. Growing economic hardship and a fall in birth rate have reduced student numbers and caused a further drop in the number of school-leavers qualifying for university entrance.

The South African government and foreign donors have attempted to revive the historically black centres by boosting their



Like all historically white institutes, the University of Cape Town (top) now has many black students.



United vision: Education minister Kader Asmal plans to merge many of South Africa's universities.

previously limited research capabilities. This year, for example, the National Research Foundation, South Africa's main research funding agency, increased its funds earmarked for these institutions by 19%. But most of the centres lack a culture of research, and some academics feel that money alone will not change this.

Geologist Gerhard von Gruenewaldt, the foundation's vice-president, says that funds are directed at fields where there is existing

E. MITSWENI/PICTURENET AFRICA; UCT

strength in research, but concedes that performance has, in certain cases, been disappointing. Others are more frank. "You don't build research capacity by throwing money at institutions," says David Jacobs, a zoologist at the University of Cape Town. "You do it by attracting good researchers to those institutions." But good researchers remain reluctant to join institutions that do not have strong reputations for research, so the government is being forced to find a different approach.

All change

After years of political procrastination, a plan to reform South Africa's higher education is now being discussed by ministers. In February, the National Working Group — a committee of business leaders, former university administrators and government officials set up in 2001 to advise on higher education — presented its report. Education minister Kader Asmal's recommendations, which are based on this report, were presented to the rest of the cabinet last month.

The plan envisages South Africa's 15 technikons and 21 universities being reduced to 21 institutions through a series of mergers. Fifteen of the 16 historically black institutions face mergers, in most cases with historically white counterparts. The new centres would comprise ten universities, seven technikons and four 'comprehensive institutions', which would offer both university and technikon programmes. The new institutions, argued the working group, would then be viable in terms of student numbers.

For institutions with a history of struggle, the plans are hard to accept. Fort Hare, for example, has a unique background to protect. Based in Alice in the southeast of the country, it was established in 1916 by white liberals and the black middle class. The school's tradition of 'plain thinking and high living' attracted black students from across the country and the rest of the continent. Its mainly white staff also included some black academics, and many contemporary South African politicians studied there. Although it has suffered from the 1960s onwards, when the white government ruled that only speakers of one ethnic language — Xhosa — could attend the school, Fort Hare is still proud of its liberal origins.

Mixed feelings

This history makes it difficult for Derrick Swartz, the university's vice-chancellor, to accept the proposed merger with nearby Rhodes University in Grahamstown and the medical school of the University of Transkei in Umtata, 400 kilometres to the east of Fort Hare. "The report fails to recognize the importance of institutional culture and identity — elements that play a crucial role in giving a particular character to universities across the world," he says.

Others say that the mergers risk damaging

islands of excellence that exist in the historically black universities. Microbiologist Don Cowan last year left University College London for the University of the Western Cape in Cape Town. Formerly reserved for mixed-race students, this university is home to a biotechnology and bioinformatics programme that is partly funded by the National Research Foundation. Cowan says that he made the move because he felt his new university was on the way up, but now fears that a proposed merger with nearby Peninsula Technikon could be damaging. "It would be a forced blending of two structures with different teaching objectives," he says.

Some academics at historically white institutions are also wary of the plans. Those at Rhodes University question the logistics of their proposed merger with Fort Hare, as the two institutions are 100 kilometres apart. And some at the University of Port Elizabeth, which is on the country's southern coast, fear that their status will be downgraded by plans to create a comprehensive institution by merging the university with the city's technikon.

But civil servants seem keen to push Asmal's recommendations through. "We cannot allow the transformation agenda to be held to ransom by those who wish to protect their own turf," wrote Thami Mseleku, director-general of the government's education department, in the South African newspaper *Business Day* on 7 May.

Joint efforts

Students who have experienced both types of institution can see both sides of the problem. Victor Rambau graduated from the historically black University of Venda, situated in a rural area near the Zimbabwean border, and is now a doctoral student in population genetics at the historically white Stellenbosch University near Cape Town. He stresses the importance to rural black communities of having undergraduate institutions nearby. "But poor administration needs to be addressed," he says. "Mergers will only work if they result in effective management procedures being put in place."

In the years that the government has taken to tackle the problem, some researchers have taken it upon themselves to improve matters. Maarten de Wit, a geologist at the University of Cape Town who is regarded by many as the country's leading Earth scientist, has initiated a collaboration with the University of Venda. Backed by funding from the mining industry, de Wit helps to supervise the research projects of three students on geology degrees at Venda.

De Wit says a lack of resources at Venda means that the students have not learnt many basic experimental skills. "During the course of their degrees, they have not used a petrologic microscope or spent anywhere near enough time doing field work," he says. The scheme has the enthusiastic support of



Connected: Maarten de Wit (above) has initiated collaborations between universities.



the Venda students, who say it has enabled them to access the vastly superior resources at the University of Cape Town.

Similar collaborations exist between other universities. But although the schemes can have a useful local impact, those involved know that the underlying issues will eventually have to be tackled at a national level.

With the government at last poised to do so, university and technikon leaders are lobbying to get the best for their institutions. Faced with conflicting demands, the government seems uncertain how to proceed. A decision on the merger plan had been due by early May, but has now been postponed until the end of the month. With no obvious solution on the horizon, the government is certain to disappoint some of those involved in the short term. But the long-term prize is a higher-education system that befits Africa's rainbow nation. ■

Michael Cherry is a zoologist at Stellenbosch University and *Nature's* South African contributing correspondent.

♦ http://education.pwv.gov.za/what's_new/what's_new.htm

Restricting genome data won't stop bioterrorism

Release of sequence information is necessary to aid defence against these pathogens.

Sir— With the “War on Terrorism” still firmly on the political agenda, measures for increasing security against possible bioterror attacks are being hotly debated. Some proposals are common-sense security issues that have long needed addressing. Others would curtail the academic freedom of investigators working on pathogenic organisms, and inhibit research into the diseases they cause (see *The Economist* **362**, 8263, 75–77; 2002, and *Nature* **414**, 237; 2001).

One particular concern is whether sequencing centres should publicly release genetic sequence information of pathogenic organisms that could be used for bioterrorism. We believe restrictions on data release would compromise the ability of legitimate scientists to study these organisms and investigate new countermeasures, while offering essentially no tangible benefit to the security of the general public.

The temptation to prohibit sequence release arises from the belief that the data will be used by hostile governments, terrorist groups or even individual bioterrorists. However, most of the problems associated with producing and using biowarfare agents — such as growth, storage and dispersal — will not be assisted by genomic information. In contrast, most defensive interventions (novel drugs, vaccines and surveillance methods) could be vastly improved by genome-based

technology. Many pathogens considered as theoretical biowarfare agents in the developed world are existing public health threats in the developing world. For example, *Yersinia pestis* still causes plague and *Bacillus anthracis* kills livestock. The release of data facilitates and accelerates research on these organisms to help developing countries that have limited resources but an immediate interest in stopping the spread of these diseases.

Prohibiting data release will not stop transmission of data. Like the organisms themselves, data can rapidly multiply and spread. If security is compromised and the data become available on the Internet, it will be impossible to track who has subsequent access to them — any party wishing to procure the information will eventually obtain it. Therefore, restricting data release implies tight containment: very few users in very few locations. Such an approach would be disadvantageous because the wider scientific community (overwhelmingly in favour of preventing hostile use of biowarfare organisms) could not then work on countermeasures.

Because few, if any, genomic sequencing facilities have the capability of handling classified information, draconian restrictions prohibiting release would restrict our capacity even to generate these data. If they are so dangerous, one would have to argue that they should not be generated at all. This is wrong. We should

continue to distribute these data freely.

Fundamentally, the primary issue of data release is control. If genome-sequence information is not available openly to the public, it is in the control of someone else, in this case government organizations. When citizens know that the government is withholding information, they become distrustful of the motives of the scientists involved in the research. If control of information about bioterror pathogen genomes is taken from the free academic community there is also the danger of a gradual encroachment on other areas of study. This arises from fears of strain ‘improvement’ to make more effective bioweapons through introduction of foreign genes. Hence, many organisms not considered biowarfare pathogens but containing toxin genes (for example *Vibrio*, *Staphylococcus* and *Salmonella*) could be restricted because they are potential sources for bioweapons.

Scientists have been branded naïve for making available sequences of agents that could be used by terrorists. But is there a difference between releasing the sequencing data for such microbial genomes and releasing the prime target for biowarfare attacks: the human genome?

Timothy D. Read*, Julian Parkhill†

*The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, Maryland 20850, USA

†The Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK

What's in a name? In the Middle East, everything

Sir— Since 1994 I have been revising the species of Israel and adjacent areas of the particular insect group on which I am a specialist. For the purposes of my entomological work, I defined five areas: Gaza Strip, West Bank, Israel, Golan Heights and Sinai, which together made up the study area, and I listed the records of the various species under these geographical headings.

I defined Israel by its pre-June 1967 borders with Jordan and Syria (the “Green Line”). I referred to the Golan Heights as being in Syria, in accordance with the UN Security Council Resolution no. 497 (1981). Finally, I considered the Gaza Strip and West Bank beyond the Green Line as occupied Palestinian territory.

Because the official “Israel Touring Map 1: 250 000” does not mark a border between Israel and the occupied or annexed territories, I had to consult

other maps to assign localities to the geographical entities. I mainly used the *Times Atlas of the World*, Comprehensive Edition (1986 and 2000), but also others, including the 1949 armistice map (available on the Internet).

I wanted to publish my manuscript in an Israeli journal because most of the material is housed by Tel-Aviv University and was collected by its staff, so I asked a member of the editorial board of the *Israel Journal of Entomology* to ask the editor if the journal would in principle accept my geographical terms. The answer was no: the journal would not use “unofficial territorial names” such as West Bank or Gaza Strip “as long as the boundaries in the Middle East have not been officially decided”. The implication was that all the localities must be assigned to Israel.

Thus the *Israel Journal of Entomology* bans any mention of Palestinian national territory and forbids the use of the two names that are used for it by the world at large. Similarly, the Saudi Arabian journal

Fauna of Saudi Arabia uses the word “Palestine” instead of “Israel”. This type of action is politics, not science.

Knut Rognes

Stavanger University College, PO Box 8002, N-4068 Stavanger, Norway

Foundation's funding

Sir— In your News story (*Nature* **417**, 3; 2002) about the new organizations formed by Craig Venter, you say “The endowment reportedly kicks off with \$100 million of Venter's own money”. To clarify, the money which will be housed in the J. C. Venter Science Foundation was from the TIGR (the Institute for Genome Research) endowment that was realized from stock Dr Venter directed to TIGR from his association with two for-profit companies.

Heather E. Kowalski

TIGR Center for the Advancement of Genomics, 1901 Research Boulevard, Rockville, Maryland 20850, USA

Science is a computer program

If complexity arises from simple rules, should we rethink how to do science?

A New Kind of Science

by Stephen Wolfram

Wolfram Media: 2002. 1,192 pp. \$44.95, £40

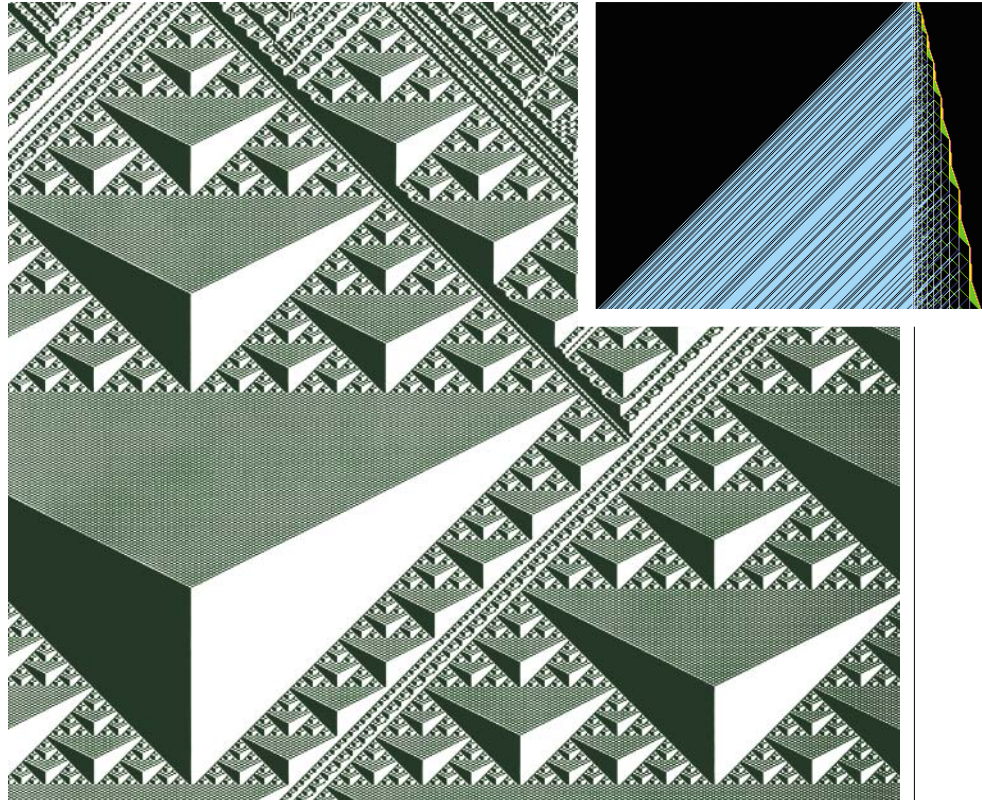
John L. Casti

This is a spectacular, iconoclastic book in almost every sense that matters—in the scope of its ideas, in its claims about science, in its sheer physical size, and in many other ways too, both good and less good. Given the book's claims and its author, it is bound to attract widespread attention and scrutiny (see, for example, *Nature* **417**, 216–218; 2002).

So let me state the volume's main thesis right at the outset. The core of Wolfram's argument, justifying the title of the book, is that the computer changes everything. Thus the primary vehicle for the scientific study of nature and humankind should now be "simple" programs and their interactions, rather than mathematical equations. The phenomena we see in the world around us should be thought of as the running of myriad simple computer programs, Wolfram argues. And the best way to understand these processes is by modelling them on a computer, not by working out the implications of an idealized mathematical model stemming from a set of equations. That's it. This is the "new" kind of science.

Before giving an example, it is important to understand what Wolfram means by a "simple program". In scientific jargon, what he means is a cellular automaton. This is simply a region, such as a line or a plane, that is divided into a large number of 'cells' like the squares on a chequerboard, each of which can be in one of several states, often just two (black or white, say). There is a rule that describes how the state of each cell changes from one moment to the next depending on the states of its neighbours. The rule of change is the 'simple program'—and there is a bewildering variety of them, even for cellular automata with just two states. The number of possible rules increases geometrically with the number of states. So once the initial configuration of the cellular automaton and the rule of change is specified, along with a definition of what counts as the 'neighbourhood' of a cell, you simply turn on the computer and let it change the state of each cell in accordance with the rule as time unfolds. The patterns made by the cells as they follow the rule of change is what Wolfram correlates with observed patterns in the real worlds of biology, chemistry, physics, economics, astronomy and all the other activities that concern scientists and humans in general.

To give a flavour of the kind of argument



Follow the rules: simple programs, or cellular automata, can produce highly complex patterns.

that pervades the entire book, think of the formation of a snowflake. Wolfram says we should start with a grid of hexagonally shaped cells in which one cell is black and the rest are white. Then, using the rule that each cell becomes black when exactly one of the cells adjacent to it was black on the previous step, a pattern emerges after about 30 steps that looks strikingly similar to a real snowflake.

Does this mean that this primitive rule of state transition is one of the rules that Nature uses to construct a real snowflake? Maybe. Or then again, maybe not. This points out a serious difficulty in using the "new kind of science", because it is easy to show that for any given snowflake pattern there are usually an impossibly large number of very different rules that will all lead to that same pattern. Which one does Nature use? Just one? Some of them? All of them? The only way to single out Nature's rule from the pretenders is to provide additional information. But the book is rather silent on how to do this. Of course, the question of model validation is a standard question in mathematical modelling, and is certainly not confined to the approach Wolfram puts forward here. But his arguments do nothing to advance this particular art.

But this is a quasi-quibble. The fact is that the book ranges over a dazzling array of

topics and areas, treating each from the very same 'simple program' point of view. The growth of crystals, fluid flow, animal pigmentation patterns, price movements in finance, cosmological models of the Universe, elementary particles, the relationship between space and time, gravity, visual and auditory perception, randomness, cryptography, consciousness, quantum theory and much, much more all make their appearance—and not just *en passant*. That's why this phenomenal book is over 1,200 pages long.

Wolfram summarizes these myriad investigations in what he terms the "Principle of Computational Equivalence". In everyday language, this principle asserts that almost all natural and artificial processes that are not obviously simple correspond to computations that are of equivalent complexity. In short, it doesn't matter how simple or complicated the rules or the initial conditions are for a process, any such process will always correspond to a computation of equivalent difficulty or, as Wolfram puts it, "equivalent sophistication". In slightly more technical terms, almost every physical and human process that can occur in our Universe corresponds to the unfolding of a program that can be run on a universal Turing machine. Note that this does not rule

out other computational architectures, such as quantum computers. But it asserts that any such non-Turing sort of computation does not correspond to any natural process in the Universe we inhabit.

The book argues that this principle is a fundamental law of Nature, with implications that go far beyond other laws such as those of thermodynamics. In fact, says Wolfram, it has “vastly richer implications... than essentially any single collection of laws in science”. This is the distilled essence of this fascinating, frustrating and overwhelmingly hubristic book: almost every process known to humankind corresponds to a computation, and every such computation is equivalent in sophistication to every other.

One of the odder aspects of this strange and wondrous book is that, although it ranges far and wide over absolutely every part of the human and natural landscape, it contains not one single reference! Wolfram’s explanation for this truly bizarre fact is that he consulted thousands of books, articles

and websites in preparing this volume (and who could really doubt that?), so it would be impossible to prepare a truly scholarly bibliography. Rather, the reader should just read the book’s 350 pages of notes, and then search on the web for a “vastly more complete picture of available references than could possibly fit in a book of manageable size”. Well, the book is already of unmanageable size, so why not spend another hundred pages and a few months to finish things off?

On a happier note, all of this material is described in entirely non-technical language, completely accessible to almost any reader. So you’ll learn a lot about many things by reading this book. In fact, it is almost like reading a small encyclopedia of science, mathematics, computing and modelling. Seen in that light, perhaps 1,200 pages is not too much after all for a tour of almost every nook and cranny between God, the Universe and just about everything else.

A New Kind of Science is a book that simply cannot be ignored. It makes spectacular

claims right from its title page onward that the author makes a herculean effort to deliver on. Whether the book’s arguments convince you or not, it will force you to reconsider your notions of what constitutes the practice and content of science. Such a book appears only once every few decades. ■

John L Casti is at the Institute of Econometrics, Operations Research and System Theory, Technical University of Vienna, A-1040 Vienna, Austria.

Optics for the real world

Classical Optics and its Applications

by Masud Mansuripur
Cambridge University Press: 2002. 512 pp.
£75, \$110 (hbk); £29.95, \$44.95 (pbk)

Chris Dainty

Many physics students today are not given the opportunity to study classical optics. This tragedy is attributable directly to their professors. What value can there possibly be in a subject, based on the concept of a light ray and the principle of diffraction, that is hundreds of years old? The professors bypassed classical optics in favour of the latest bit of trendy physics that could earn them their tenure. How many professors have any idea how to design a lens or do a simple exercise in classical optics? This task is too trivial, isn’t it?

This book is a series of 37 vignettes, adapted from articles written for *Optics and Photonics News*, a monthly magazine produced by the Optical Society of America. Masud Mansuripur describes phenomena that we all know we should understand, even if our recollection is hazy. The topics are classics: Abbe’s sine condition, gaussian beam optics, evanescent coupling, ellipsometry... these are fundamental phenomena that have practical application in almost every product that uses optics. Forget photonics, it is classical optics that turns today’s optical inventions into real products and processes that benefit society.

Throughout the book Mansuripur uses computer calculations and simulations to illustrate each topic. On balance, this adds considerably to the clarity of the descriptions, and is an aid to understanding. On the other hand, life is real, not virtual, and much of optical engineering and applied optics is devoted to achieving in practice what any computer program can demonstrate in a fraction of a second. I would have liked to see more comparisons of the real and virtual worlds. The issue here is that simulations rarely model the imperfections of the real world properly.

One of the positive aspects of the book is that it encourages the reader to delve deeper



Starry eyed: as this engraving from 1820 shows, the use of optical devices has a long history.

into each topic. The author writes in a relaxed style and manages to make each subject sound easy. The computer simulations appear to be done effortlessly.

As I read the chapters, I felt I was learning a lot. But when I came to subjects that were familiar to me (such as the van Cittert–Zernike theorem, or the Shack–Hartmann sensor), I realized that I would have presented the subject differently, probably placing more emphasis on the fundamentals. Partial coherence of light is fundamentally about the statistical nature of light, but this is not really emphasized here, even though it can be illustrated graphically using computer-generated examples. I suspect that there is some non-uniformity in the coverage of each topic.

Even so, I thoroughly recommend this book to anyone in optics who is interested in doing something useful. Everyone will learn something, and refresh their memory on subjects that are fundamental to so many practical devices that use optics.

Chris Dainty is in the Optics Section, Blackett Laboratory, Imperial College, London SW7 2BZ, UK.

Swimming lessons

Molecular Biology of the Cell, 4th edition

by Bruce Alberts, Alexander Johnson, Julian Lewis, Martin Raff, Keith Roberts & Peter Walter

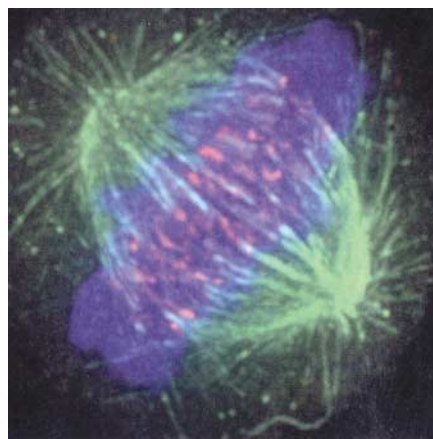
Garland Science: 2002. 1,400 pp. £75, \$102 (hbk); £44 (pbk)

Angus I. Lamond

A generation of students have learned the basics of molecular cell biology thanks in no small part to courses based on the pioneering textbook *Molecular Biology of the Cell*, which was first published 18 years ago. Through three editions it has established an enviable record for high-quality presentation, with the authors showing a remarkable ability to make both basic concepts and cutting-edge research topics accessible to readers. The arrival of the fourth edition prompts the obvious question: "What's new?"

Not surprisingly, size is the first thing. The new edition is even larger than its predecessors, reflecting the vigorous activity of the field and the inexorable expansion of detailed information regarding cellular processes and molecular structures and interactions. In an era when we are drowning in information, the job of a textbook such as *Molecular Biology of the Cell* is to help teach its readers how to swim. This it does admirably, and the punctilious attention to detail and effort devoted by the authors to covering this huge field in a lucid and easy-to-read style shines through on every page.

The material is now presented in five sections, with the one previously entitled



Seeing cells: composite stained micrographs can reveal the cell's activities, such as mitosis.

'Molecular Genetics' now being divided into 'Basic genetic mechanisms' and 'Methods'. The latter section highlights the importance of experimental approaches to cell biology and explains how most of the key findings discussed elsewhere in the text were made. This is something I try to emphasize when I lecture to students. The section is not a guide for a practical course, nor does it provide laboratory protocols; rather, it explains the basic theory behind a wide range of commonly used molecular and microscopy techniques.

The descriptions are generally clear and well illustrated, although I noticed a few rare exceptions. For example, fluorescence resonance energy transfer is poorly explained and oversimplified to the point of being quite misleading. The now widely used application of proteins tagged with green fluorescent protein for photobleaching studies is a surprising omission, and in several cases, such as the discussion of two-dimensional gel electrophoresis, I feel that some comment on the limitations of the technique, as well as its advantages and applications, would have offered a more balanced view. But overall, the methods section is comprehensive, well written and a genuine improvement on the previous edition.

The availability of complete genome sequences for many model organisms is an obvious new topic for the fourth edition, and material has been added to many chapters to incorporate this information. Major advances in understanding specific processes, such as nucleo-cytoplasmic protein transport, histone modification and chromatin remodelling, are all included. A new chapter on the exciting advances in studies of innate immunity and interactions between host cells and pathogens is a particularly timely and welcome addition.

With such a wide range of subjects being covered, I am impressed at how well the information has been incorporated into a house style. In most cases the topics are summed up clearly and succinctly, making good use of illustrations to explain each

point or concept. Students can use many of the chapters directly as cram notes for exams, and I expect that many examples of prose from the text will be appearing verbatim in future essays. Although this is a generally successful formula, it doesn't work well for some topics where the field has simply not advanced far enough to lend itself to a short and simple explanation.

I do wonder, however, whether *Molecular Biology of the Cell* has now outgrown its present format. Its sheer physical size renders it somewhat unwieldy and I found myself spending a lot of time flicking between chapters and from one section to another to follow up on linked topics. More extensive cross-referencing to help the reader find related material would have been useful, but ultimately there is no way to avoid this dispersal of information in a printed textbook. As a result, I kept wanting to 'click' on highlighted keywords that would provide these cross-references and links between topics and sections.

The material is crying out for presentation in an electronic format. I therefore turned to *Cell Biology Interactive*, the new CD-ROM that accompanies the fourth edition, to see if it would add this missing ingredient. Unfortunately, it does not. Unlike the comprehensive presentation and uniform style of the printed text, I found the CD patchy and variable, both in terms of the topics covered and the quality of the material provided. Nonetheless, there are some very useful items on the CD, and the ability to view and rotate three-dimensional protein and nucleic-acid structures and to see time-lapse movies of cellular processes, including cell movement, cytokinesis and gastrulation, will undoubtedly be helpful to students.

Overall, the CD convinced me that an interactive format, rather than increasingly bloated printed volumes, is the best way to present molecular cell biology to students in the future. But in my opinion, *Cell Biology Interactive* is more of an 'add-on' than an integral part of the fourth edition. It scratches the surface of what is now possible with this format and fails to live up to the level of innovation that we have come to associate with the book.

In conclusion, the fourth edition of *Molecular Biology of the Cell* is an outstanding textbook of the sort I would have liked to have when I was an undergraduate. I will certainly recommend it to my own students. The clear and concise style and copious informative diagrams and illustrations set an impressive standard, and the comprehensive coverage of the field that has been achieved is quite remarkable. I hope that a future fifth edition will not only maintain this standard, but will do more to take advantage of new technologies and interactive material.

Angus I. Lamond is at the Wellcome Trust Biocentre, University of Dundee, Dundee DD1 5EH, UK.

In the balance

Ivar Ekeland

One could write a history of science around the concept of equilibrium. It originates in mechanics, denoting a situation in which a balance of conflicting forces results in rest. It moves into physics as thermodynamic equilibrium, a situation in which a flurry of microscopic motions results in macroscopic rest. It then crosses a new boundary, entering the social sciences as economic equilibrium, a set of prices that equate supply and demand across goods, thereby enabling producers and consumers to trade.

During this journey, the mechanical analogy had always served as a trustworthy guide — in, for example, such expressions as 'market forces'. Yet the concept has now taken on a new meaning, and biological or sociological analogies have become much more relevant. Game theory, starting with the seminal work of John von Neumann and John Nash, has had its main effect in the study of systems in which decision-making

agents interact in a complex way, recrafting equilibrium into a coherent, powerful and widely applicable concept. In the social sciences, an equilibrium is a situation that involves several individuals or groups, in which each one's actions turn out to be the best reply to everyone else's. It is a situation of stable mutual adjustment — everyone anticipates everyone else's behaviour, and all these anticipations turn out to be correct. In other words, it is a set of self-fulfilling prophecies that individuals formulate about each other's actions.

Such situations are central to social life because they are the only stable ones. Not to be in equilibrium means that some anticipations turn out to be wrong, so that some actions are inappropriate in a particular situation. This leads the affected individuals or groups to revise their anticipations and adjust their actions, thereby creating new discrepancies to be corrected at the next stage. Hence the entire situation is destabilized, and the system begins to oscillate wildly. In equilibrium, on the other hand, all anticipations are confirmed by experience and every acquired behaviour turns out to be appropriate to every situation, so that anticipations become more ingrained as time goes on, eventually solidifying into social norms.

Basic features of social organization, such as trust and power, simply express some underlying equilibrium. Power is nothing but the illusion of power, the universally held belief that a certain person is to be obeyed, that certain orders are to be followed. It is self-fulfilling, for if I am given an order by such a person, I will follow it for the simple reason that if I don't, someone else will, and it will probably be the worse for me. Trust is the belief that others will comply with certain rules; each time I myself comply with those rules I strengthen the general feeling of trust.

Note that distrust is also self-supporting. If I distrust you and you distrust me, I will take precautions to protect myself against your anticipated behaviour, giving you good reason to distrust me a little bit more; sadly, we are witnessing precisely this situation in the Middle East. Trust is an equilibrium, distrust is another. In the former, everyone trusts everyone else, and is right to do so; in the latter, everyone distrusts everyone else, and is right to do so. A situation in which some trust but others don't would not be stable, for the first group would soon learn to join the second. The main (unanswered) question of social science is thus: which equilibrium will prevail?

Reaching equilibrium does not even require intelligence — a darwinian process without conscious decisions will do just

Equilibrium

In the social sciences, an equilibrium is a situation that involves several individuals or groups, in which each one's actions turn out to be the best reply to everyone else's.

as well. Natural selection leads competing species to an equilibrium — a stable adjustment to each other. Off the west coast of South Africa, the waters of Malgas Island are dominated by seaweed and rock lobsters that prey on mussels and whelks. Nearby Marcus Island is similar in every respect, but its waters have extensive mussel beds and whelks at high density, lobsters and seaweed being notably absent. In a famous experiment, Amos Barkai and Christopher McQuaid transferred thousands of adult lobsters from Malgas to Marcus — the whelks simply ate the much bigger creatures, and within a week there was not a single lobster left. Yet local fishermen claim that there were lobsters on both islands around 20 years before.

The prevailing equilibrium on these islands is therefore not determined by a local condition but by past history — a catastrophic event must have occurred to switch Marcus to another equilibrium. This example shows how rules that we assume to be universal, such as predator-prey relationships, in fact are relative to some equilibrium. If the lobsters on Malgas were able to think, they would believe it to be a rule of nature that lobsters prey on whelks, whereas just the opposite is true one island away.

Humans are also susceptible to this illusion: we are born into an equilibrium, the extent of which we do not appreciate. Our ethics of behaviour, our sense of justice, our social habits and our individual rights are not reflections of absolute truths or of human nature. They are relative to an equilibrium that our species, or society, has reached. Social norms and mores have no reality other than our belief that others behave according to them — they are self-fulfilling prophecies. Our social fabric and our personal ethics, everything we stand for, rest on equilibria, which we conjure out of nothing. We are such stuff as dreams are made on. ■

Ivar Ekeland is at the Institute of Finance, Université Paris-Dauphine, 75775 Paris, France.

FURTHER READING

Barkai, A. & McQuaid, C. *Science* **242**, 62–65 (1988).
Sigmund, K. *Games of Life: Explorations in Ecology, Evolution and Behaviour* (Penguin, London, 1993).
Crozier, M. & Friedberg, E. *L'acteur et le Système* (Seuil, Paris, 1983; English transl. ISBN 0608092924).



Balancing act: in society, as in acrobatics, a single wrong move can upset the fragile equilibrium.

Anti-malarial mosquitoes?

Gareth J. Lycett and Fotis C. Kafatos

Researchers have generated genetically altered mosquitoes with a reduced ability to transmit the malaria parasite in the lab. This is an important proof of principle, but there's still a long road ahead.

Ever since molecular biologists discovered how to genetically transform fruit-flies, malaria researchers have dreamt about using similar techniques to turn disease-carrying mosquitoes into harmless pests. If such 'transgenic' mosquitoes were genetically fit but unable to transmit the malaria-causing parasite, *Plasmodium falciparum*, and if they could be introduced into natural populations at a high enough frequency, they might reduce the prevalence of this disease. Small but significant steps have been taken towards this aim in recent years. A major breakthrough was the development of methods for introducing desirable genes into mosquitoes. And genes that confer anti-malarial properties have been identified. On page 452 of this issue, Jacobs-Lorena and colleagues¹ describe how they have combined these advances to create a transgenic mosquito that transmits a *Plasmodium* parasite at much diminished rates in the laboratory.

Malaria is a devastating scourge. Recent estimates are that every year, chiefly in Africa, the malaria parasite infects 300–500 million people and causes 1 million to 3 million deaths², as well as lowering economic growth by 1.3% of GDP. Only a few mosquito species transmit the parasite, and in the past, control of these carriers by massive spraying of insecticides³ or eliminating breeding sites has greatly limited malaria in some areas — even eradicating it in a few places, such as southern Europe after the Second World War. But the toll from malaria is growing. Mosquitoes develop resistance to insecticides; the parasites become resistant to preventative and therapeutic drugs; effective vaccines do not exist; and poverty, together with social and ecological disruption, hinders the application of current control measures. This disease is too complex to yield to a 'magic-bullet' solution, but new mosquito-based methods of control will certainly help.

The malaria parasite has a complicated life cycle (Fig. 1). After a mosquito feeds on an infected human, taking up parasite-laden blood, the parasite develops into a motile form — the ookinete — in the cavity of the mosquito's midgut. The ookinete then traverses the midgut wall (its epithelium) and, on the wall's outer face, transforms into a structure called an oocyst, in which thousands of worm-like sporozoites develop. In time the sporozoites burst out into the open blood space (the haemocoel), later reaching

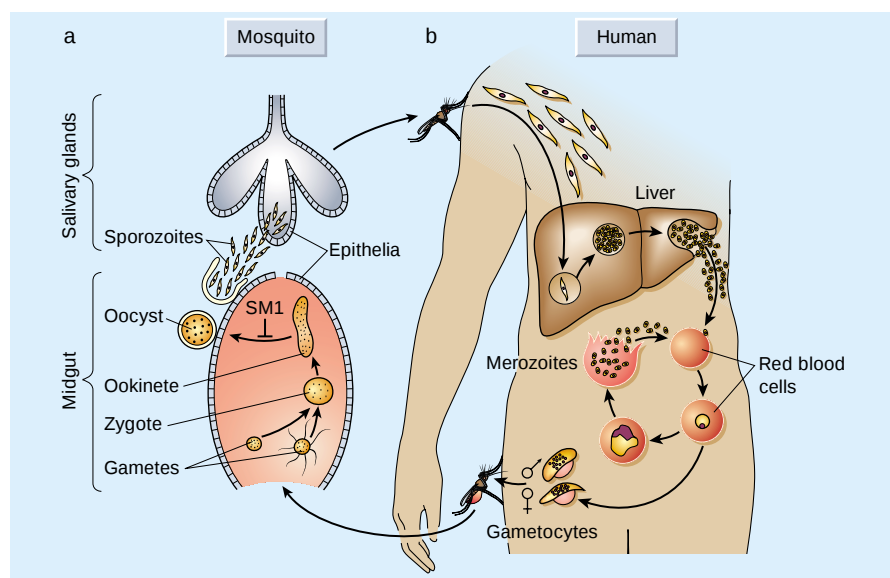


Figure 1 Life cycles of the malaria parasite. **a**, In mosquitoes; **b**, in humans. **a**, The parasite's life cycle in the mosquito begins when the mosquito bites an infected human, taking up blood that is laden with the precursors of male and female parasite gametes (gametocytes). These fuse in the insect's midgut to form a zygote, which then develops into the motile ookinete. The ookinete crosses the midgut epithelium and forms an oocyst, filled with sporozoites, on the outer surface of the epithelium. Later the oocyst bursts and the sporozoites travel to the salivary glands. **b**, When the mosquito bites another human it injects the sporozoites, which journey to the liver, there developing into merozoites, which can multiply in red blood cells. A fraction of the merozoites develop into gametocytes, ready to be taken up by the mosquito. Jacobs-Lorena and colleagues¹ have engineered mosquitoes to express the peptide SM1, which seems to interfere with how the parasite detects mosquito epithelia.

and invading the epithelia of the salivary glands. When the mosquito feeds again it spits out the sporozoites, which start the human phase of their life cycle by infecting another person.

In theory, then, it should be possible to hinder transmission of the parasite by interfering with how it binds to and traverses the mosquito midgut epithelium. Researchers are currently using molecular and cellular biology to study these events^{4,5}. The parasite is thought to recognize the midgut epithelium by detecting particular epithelial proteins, and in earlier work Jacobs-Lorena and colleagues⁶ took a short cut to identifying these 'receptors'.

The authors⁶ screened large numbers of random artificial peptides (short amino-acid chains) and fished out one, termed SM1, that bound to mosquito gut and salivary glands but not other tissues. This suggested that the peptide bound to a molecule, possibly a protein, that is found only on the gut and salivary glands. Could this molecule be the hypothet-

ical receptor used by the parasite to identify target tissues? Support for this idea came from the finding that the SM1 peptide inhibited parasite invasion when injected or fed to mosquitoes — presumably because it blocked the receptor, stopping the parasite from getting a hold on the epithelium. It was predicted that a transgenic mosquito that produced this peptide in the gut or salivary glands would be inhospitable to the parasite.

There have been previous attempts to genetically engineer mosquitoes that are resistant to the malaria parasite. For instance, De Lara Capurro *et al.*⁷ used a virus to make the salivary glands of *Aedes aegypti* mosquitoes express antibodies against a protein on the parasite surface. This interfered with binding of the parasite to the salivary glands but, because of the experimental design, not in a heritable way. Once reliable techniques for stable (heritable) genetic transformation of mosquitoes were discovered^{8,9}, Kokoza *et al.*¹⁰ genetically enhanced the immune system of *A. aegypti* by forcing

them — after blood-feeding — to over-express an introduced gene encoding the antibacterial peptide defensin. Extracts from these mosquitoes killed bacteria *in vitro*. But anti-parasite activity was not reported.

Now, Jacobs-Lorena's group¹ has succeeded in generating a 'strain' of *Anopheles stephensi* mosquitoes that carry an inherited piece of DNA encoding both SM1 and a 'signal' peptide, which helps transport SM1 into the gut cavity. The regulatory region of this artificial gene induces the production of SM1 when needed — after blood-feeding. The authors then allowed transformed and untransformed mosquitoes to feed on the same parasite-infected mouse. In nine experiments the number of transformed mosquitoes carrying oocysts was reduced by 47% on average, and the number of oocysts per mosquito dropped by 80%, relative to untransformed insects. The authors also allowed three groups of transformed mosquitoes to develop for 25 days after feeding, and observed sporozoites at 1.7–19% of control levels. Two of these groups could not transmit the infection to mice. The third showed a 50% reduction in transmission. As the SM1-coding DNA is stably inherited by the mosquitoes' offspring, Jacobs-Lorena and colleagues have shown the feasibility of generating populations of transgenic mosquitoes that have diminished potential to carry the malaria parasite.

This is a proof of principle and as such is a milestone in malaria research. Still, molecular biologists who study mosquitoes fully appreciate the length of the road ahead¹¹. Jacobs-Lorena and colleagues¹ used a parasite that causes malaria in rodents, and it remains to be seen whether the SM1 peptide is effective against human parasites carried by other mosquito species. Moreover, to ensure it will be effective in a given location, transformation would probably need to be done in wild populations taken recently from local sites, rather than laboratory strains. There is also no firmly established method for driving transmission-blocking genes through mosquito populations. Too little is known about natural populations and gene flow between mosquito subspecies to allow us to predict the fate of introduced genes. The consensus in the mosquito-research community is strongly against premature field experiments. It is felt that even fully contained field trials must await stringent laboratory experiments and long-term population studies, and that transformed mosquitoes should meet the requirement of a 'significant probability' of reducing malaria prevalence before being released.

The new work¹ is exciting, nonetheless, and represents a new era of malaria-related research. 'Reverse' genetic analysis, in which mutated forms of normal genes are introduced back into an organism to study their function, is becoming routine in *Plasmodium*

and no doubt will soon be common in mosquitoes. These techniques will capitalize on the data generated by genome-sequencing projects. In the past few months the raw genome sequence of the most important malarial carrier, *Anopheles gambiae*, became available on public databases^{12,13}. Expectations are that by midsummer, the annotated gene content of all three organisms involved in transmitting malaria (humans^{14,15}, parasites^{16,17} and mosquitoes) will be at hand. Thoughtful exploitation of this information should accelerate progress towards new ways of controlling malaria. These will range from new diagnostic tools, effective vaccines that protect people from the parasite and the discovery of new drug targets in the parasite itself, to transmission-blocking vaccines and even mosquitoes that are rendered ineffective as carriers. ■

Gareth J. Lycett and Fotis C. Kafatos are in the European Molecular Biology Laboratory, Meyerhofstrasse 1, 69117 Heidelberg, Germany.

Polymer chemistry

Swell gels

Jindřich Kopeček

Linked chains of polymers can form hydrogels, whose properties are attractive for biomedical applications. It seems that the molecular arrangement of the polymer ingredients is central to hydrogel performance.

Hydrogels are three-dimensional networks of polymer chains that swell, but don't dissolve, in water. The formation of hydrogels is an interesting phenomenon, and could one day supplement tools for the design, synthesis and self-assembly of novel biomaterials and drug-delivery systems. On page 424 of this issue, Nowak *et al.*¹ investigate the hydrogels formed by a family of polypeptides that would not usually be expected to adopt such a structure. The hydrogels formed are relatively robust, up to temperatures of around 90 °C, and rearrange rapidly to recover their structure after experiencing stress.

Nowak *et al.* studied diblock copoly-peptide amphiphiles — polymers composed of charged (hydrophilic) and hydrophobic blocks of amino-acid residues. In solution, diblock copolymers — which are made of two chemical building blocks, say, A and B (Fig. 1a) — generally form 'micelles'²: if block A is soluble in water but block B isn't, in aqueous solution a hydrophobic core of B blocks forms, surrounded by a hydrophilic corona of A blocks (Fig. 1b). But instead of forming micelles, the amphiphiles studied by Nowak *et al.* linked up at low concentrations to form the three-dimensional structure of hydrogels (Fig. 1c). Moreover, this behaviour seems particular to diblock copolymers: random copolymers (Fig. 1a) did not form hydrogels.

e-mails: lycett@embl-heidelberg.de;

dg-office@embl-heidelberg.de

1. Ito, J., Ghosh, A., Moreira, L. A., Wimmer, E. A. & Jacobs-Lorena, M. *Nature* **417**, 452–455 (2002).
2. Bremen, J. *Am. J. Trop. Med. Hyg.* **64** (Suppl.), 1–11 (2001).
3. Curtis, C. F. & Lines, J. D. *Parasitol. Today* **16**, 119–121 (2000).
4. Zieler, H. & Dvorak, J. A. *Proc. Natl Acad. Sci. USA* **97**, 11516–11521 (2000).
5. Han, Y. S., Thompson, J., Kafatos, F. C. & Barillas-Mury, C. *EMBO J.* **19**, 6030–6040 (2000).
6. Ghosh, A. K., Ribolla, P. E. M. & Jacobs-Lorena, M. *Proc. Natl Acad. Sci. USA* **98**, 13278–13281 (2001).
7. De Lara Capurro, M. *et al. Am. J. Trop. Med. Hyg.* **62**, 427–433 (2000).
8. Jasinskiene, N. *et al. Proc. Natl Acad. Sci. USA* **95**, 3743–3747 (1998).
9. Catteruccia, F. *et al. Nature* **405**, 959–962 (2000).
10. Kokoza, V. *et al. Proc. Natl Acad. Sci. USA* **97**, 9144–9149 (2000).
11. James, A. A. *et al. Parasitologia* **41**, 461–471 (1999).
12. http://www.ensembl.org/Anopheles_gambiae
13. http://www.ncbi.nlm.nih.gov:80/cgi-bin/Entrez/map_search?chr=agambiae.inf
14. International Human Genome Sequencing Consortium *Nature* **409**, 860–921 (2001).
15. Venter, J. C. *et al. Science* **291**, 1304–1351 (2001).
16. <http://PlasmoDB.org>
17. <http://www.ncbi.nlm.nih.gov/Malaria>

The history of hydrogels goes back to the late 1950s when Wichterle and Lím³ synthesized hydrogels based on copolymers of 2-hydroxyethyl methacrylate with ethylene dimethacrylate — the first biomaterials designed for medical applications. These hydrogels were the original materials used for soft contact lenses, their biocompatibility having been proven during their clinical use as implant materials⁴. The commercial success of soft contact lenses stimulated enormous interest in hydrogels, and eventually led to the development of 'smart' hydrogels that change their properties after exposure to an external stimulus, such as pH, temperature, light or electric field^{5–8}.

But there are factors that limit the broad application of hydrogels, such as their relatively long response time⁹ or a large difference between the time taken to turn 'on' in response to stimuli and the time taken to turn 'off' again. For example, a swollen hydrogel might shrink quickly after exposure to a stimulus, but the re-swelling after reversal of the stimulus would happen much more slowly.

Some of these limitations are consequences of how hydrogels are synthesized. Traditional synthetic pathways, crosslinking copolymerization and crosslinking of polymer precursors, do not permit exact control of chain length, sequence or three-dimen-

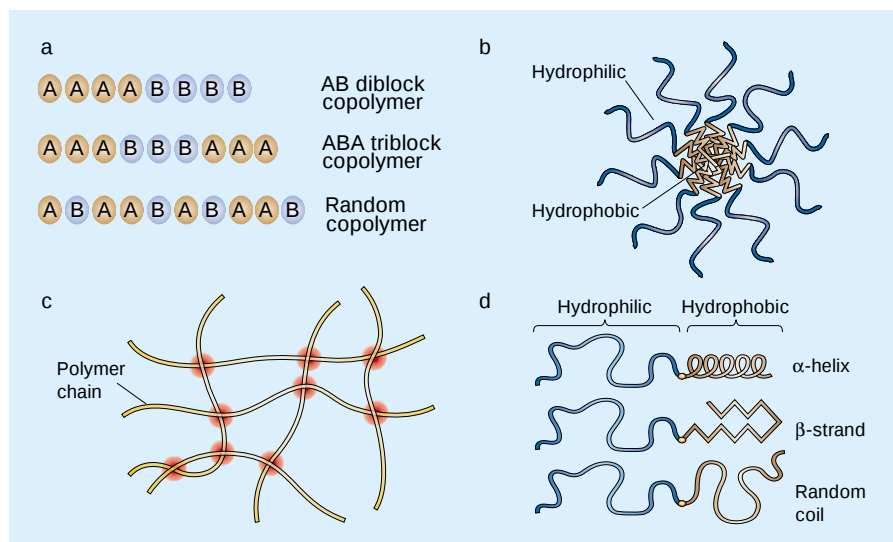


Figure 1 Copolymers, micelles and hydrogels. a, The arrangement of monomer units (A and B) in copolymer chains can take diblock, triblock or random form, and subtly affects the behaviour of the molecules. b, Amphiphilic diblock copolymers have a hydrophobic part and a hydrophilic part. In water, the molecules form structures known as micelles, with the hydrophobic blocks of each molecule forming a core surrounded by a corona of hydrophilic blocks. c, Hydrogels are three-dimensional structures of crosslinked polymer chains. Nowak *et al.*¹ have shown that some amphiphilic molecules form hydrogels instead of micelles; how easily they do so depends on the shape of the hydrophobic part of the molecule. d, According to Nowak *et al.*, molecules whose hydrophobic part is an α -helix are better at forming hydrogels than those with the β -strand conformation or a random coil.

sional structure. But two new models for the design and synthesis of smart hydrogels have emerged: the synthesis of hydrogels (or their building blocks) by genetic engineering methods and the design of associative building blocks, which self-assemble into hydrogel structures. The combination of these approaches has the potential to be a rational pathway for the design of hydrogels with a rapid on–off response.

Several groups are working on the self-assembly of block copolymers. For example, Tsitsilianis *et al.*¹⁰ have demonstrated hydrogel formation from triblock copolymers (Fig. 1a), and Petka *et al.*¹¹ have produced triblock copolymers by protein engineering that self-assemble into stimuli-sensitive hydrogels. Diblock copolymers of polyoxyethylene and polyoxybutylene have also been shown to form hydrogel structures, but only at high polymer concentrations¹².

Nowak *et al.*¹ designed and synthesized block copolypeptides with poly-(L-lysine) or poly-(L-glutamic acid) as the hydrophilic block and poly-(L-leucine), poly-(L-valine) or poly-(D/L-leucine) as the hydrophobic block. The authors show that these low-molecular-weight diblock copolymers associate into hydrogels at very low polymer concentrations. In particular, the hydrogels maintain their mechanical strength at high temperatures and recover (rearrange) rapidly after stress.

The unique properties of these hydrogels depend on their structural parameters. Nowak *et al.* show that the shape of the polymer chains is an important factor in the

hydrogel self-assembly — or gelation — process. They found that α -helical segments are better gelators than β -strands, which in turn are better than random coils (Fig. 1d). Earlier studies of superabsorbing hydrogels¹³ have shown that the greater the porosity of a hydrogel, the faster it swells. Nowak *et al.* believe that the fast response times of their hydrogels are also tied in with porosity, as the diblock copolymers were found to separate into a gel matrix and polymer-free liquid domains. They also think that the relatively low molecular weight and narrow molecular-weight distribution of copolymers contribute to the rapid rearrangement of polymer chains and fast recovery of hydrogel structures after stimulus-driven dissociation of polymer chains.

The porosity and fast response of these hydrogels may prove useful in biomedical applications, particularly for drug delivery: molecules, even large molecules such as proteins and DNA, could be loaded into the hydrogel structure to be released as the hydrogel responds to a physiological trigger. Moreover, the degradability of amino-acid-based hydrogel structures *in vivo* is attractive for biomedical applications. Other applications might include the formation of scaffolds for trapping and growing cells for tissue regeneration. One could even imagine that the self-assembly properties of the copolymers studied by Nowak *et al.* could be exploited to develop these materials into sensor/actuator systems, such as controllable membrane-separation systems and electronically controlled drug-delivery systems¹².



100 YEARS AGO

Sun-pillar(?) Miss Herschel (a careful observer) has just called me out to see one. At 7.10 p.m. she saw the sun above a bank of clouds, in a somewhat hazy sky, but no clouds above it for a space of some 5 degrees. Above that was a light-fringed belt of clouds of great depth. From the sun a parallel-sided pillar of light, just like the reflection of the sun in a slightly rippled sea, stood upright into, and stopped at, the light-fringe; it was not so bright as the reflection spoken of would be, but markedly brighter than the background sky; colour yellow. Miss Herschel had to bicycle home three-quarters of a mile uphill to call me, and it was fading before she reached home. I was prompt, but too late (7.25) to get a good view. W. J. Herschel
From *Nature* 22 May 1902.

50 YEARS AGO

The issue of *L'Astronomie* for December 1951 contains an address by M. Jean Cabannes, president of the Société Astronomique de France, delivered at a meeting on November 18, the title being "Le Ciel Nocturne". In this address an excellent account is given of the night sky, with eight illustrations, most of which are of spectra of the night sky under various conditions: in the visible part of the spectrum, 4000–4900 Å.; in the ultra-violet; and in the infra-red, 8300–11000 Å., including two of the OH-bands, and others. In spite of our knowledge of the actual atoms and molecules which cause the luminosity in the upper atmosphere, it is admitted that many problems of the phenomenon still remain unresolved. If the luminous layer is thin, what is its height, and if thick, how are the luminous centres distributed in it? If the atmosphere were homogeneous, there would not be much difficulty in determining these altitudes; but as it is heterogeneous, difficulties arise from the necessity of making a long series of measurements to give an acceptable interpretation to the mean. M. Cabannes discusses the possibility of deriving results from rockets equipped with photometers specially adapted for such observations; although the problems are still far from solved, nevertheless the number of astronomers, physicists and chemists who are devoting their attention to the subject is continually increasing, and the successes hitherto obtained give great hopes for the future.

From *Nature* 24 May 1952.

Although these are long-term perspectives, the work of Nowak *et al.*¹ has enhanced our understanding of block copolymers and their assembly into hydrogels. The new information on the relationship between the structure of polypeptide amphiphiles and their properties as hydrogels could one day form a useful basis for the design of new smart biomaterials.

Jindrich Kopeček is in the Department of Pharmaceutics and Pharmaceutical Chemistry, and the Department of Bioengineering, University of Utah, Salt Lake City, Utah 84112, USA.
e-mail: jindrich.kopecek@m.cc.utah.edu

1. Nowak, A. P. *et al.* *Nature* **417**, 424–428 (2002).

2. Moffitt, M., Khougaz, K. & Eisenberg, A. *Acc. Chem. Res.* **29**, 95–102 (1996).
3. Wichterle, O. & Lím, D. *Nature* **185**, 117–118 (1960).
4. Voldřich, Z., Tománek, Z., Vacík, J. & Kopeček, J. *J. Biomed. Mater. Res.* **9**, 675–685 (1975).
5. Kopeček, J., Vacík, J. & Lím, D. *J. Polym. Sci. A-1* **9**, 2801–2815 (1971).
6. Wang, C., Stewart, R. J. & Kopeček, J. *Nature* **397**, 417–420 (1999).
7. Suzuki, A. & Tanaka, T. *Nature* **346**, 345–347 (1990).
8. Kwon, I. C., Bae, Y. H. & Kim, S. W. *Nature* **354**, 291–293 (1991).
9. Kajiwara, K. & Ross-Murphy, S. B. *Nature* **355**, 208–209 (1992).
10. Tsitsilianis, C., Iliopoulos, I. & Ducouret, G. *Macromolecules* **33**, 2936–2943 (2000).
11. Petka, W. A., Harden, J. L., McGrath, K. P., Wirtz, D. & Tirrell, D. A. *Science* **281**, 389–392 (1998).
12. Hamley, I. W. *et al.* *Langmuir* **16**, 2508–2514 (2000).
13. Chen, J., Park, H. & Park, K. J. *Biomed. Mater. Res.* **44**, 53–62 (1999).

Circadian rhythms

A gut feeling for time

Michael H. Hastings

Many body functions keep a daily rhythm, maintained by a central clock in the brain. But how does the clock communicate with the rest of the body? The small protein prokineticin 2 looks well placed to be the messenger.

Have you ever wondered how you can wake up automatically before the alarm clock rings? It is because your body's neural and physiological preparations for wakefulness are not just a response to the world outside, but are strictly controlled by an internal clock — a region in the hypothalamus of the brain called the suprachiasmatic nucleus (SCN)¹. Neurons in the SCN have the remarkable ability to autonomously generate a cycle of electrical activity with a period very close to 24 hours. But how might this relatively tiny population of some 10,000 neurons organize the daily (circadian) rhythms of the whole body and keep us in tune with the world? On page 405 of this issue Cheng and colleagues² provide a possible answer.

A compelling picture of the core SCN clockwork has been established by molecular genetics. In essence, the canonical clock genes *Period* and *Cryptochrome* make up a self-sustaining circadian oscillator¹, in which the critical mechanism is delayed negative feedback. These genes are switched on by the proteins *Clock* and *Bmal*, and are periodically switched off by a complex of their own encoded proteins, *Per* and *Cry*. So gene turn-off inevitably follows gene turn-on, in an inexorable daily loop, and mutations that affect the loop's stability are tightly linked to inherited human sleep disorders³.

A complementary molecular-genetics approach to understanding sleep comes from the study of the gene encoding the hypocretin/orexin peptides. Mutations in this gene, or in that for the receptor protein that enables neurons to detect the peptides, are linked to the sleep disorder narcolepsy⁴, in which people uncontrollably fall into brief

periods of deep sleep. The cells that make hypocretins/orexins lie in the dorsal hypothalamus and have a powerful excitatory effect on neural systems that sustain wakefulness. But what sits between the oscillatory molecular cycle of the SCN and the neural machinery that directly causes sleep and wakefulness? In other words, what is the messenger of circadian time?

SCN neurons 'fire' in a circadian pattern, suggesting that they signal time by means of conventional electrochemical communication with other neurons. Yet their physical connections with other neurons are sparse and principally local. In fact, intracerebral transplant studies have shown that physical contact between SCN and target neurons is not necessary for them to communicate⁵. Rather, some molecule, presumably secreted from the SCN under the influence of electrical firing, confers circadian control of sleep and wakefulness.

Cheng *et al.*² now reveal that a small protein, prokineticin 2, may well be time's messenger. Consisting of 81 amino acids and identified previously as a regulator of gastrointestinal movements⁶, this protein fulfils many of the criteria expected of the missing temporal link. It shows high-amplitude, rhythmic expression in the SCN in mice, conferred by regulatory sequences within its encoding gene that are a direct target for positive and negative components of the core clockwork (*Clock*, *Bmal*, *Per* and *Cry*). It is also downregulated in mice with genetically defunct circadian clocks. As with the *Period* gene, the expression of prokineticin 2 in the SCN is activated by illumination of the retina. So, following a change in the timing of lights on and lights off, the pattern of expression of

prokineticin 2 shifts, and so does the activity/rest cycle of mice. Finally, secreted prokineticin 2 is resistant to degradation by protein-cleaving enzymes, making it ideal for long-range signalling in space and time.

How does prokineticin 2 affect behaviour? To find out, Cheng *et al.* injected the protein into the brains of rats. They found that clock time *per se* was not affected, but that the animals became much less active during circadian night — their usual active time — indicating that prokineticin 2 suppresses wakefulness. So prokineticin 2 is not a core clock component, but it does appear to communicate clock time to other brain regions. When functioning normally, the clock would ensure that prokineticin 2 levels peak during circadian daytime, and prokineticin 2 would presumably suppress the wakefulness of rats at that time.

Intriguingly, during the circadian daytime following the nocturnal injection of prokineticin 2 there was a spontaneous rebound, and the rats were active at a time when they should be asleep². This may reflect the fact that the targets of prokineticin 2 in the brain became 'desensitized' during the night because of the excess injected protein, and so failed to respond to the normal protein the next day. It is clearly important to identify these targets, and this is the final piece of news. The receptor protein that detects prokineticin 2 is expressed in brain regions that constitute a roll call of neuro-anatomically defined SCN targets, including the dorsal hypothalamus with its hypocretin-releasing cells. Through these targets, SCN-derived prokineticin 2 can signal circadian daytime to several downstream systems, coordinating circadian behaviour and physiology.

The receptor is also expressed abundantly in the SCN, intimating that it may have other roles. Indeed, Cheng *et al.* show that prokineticin 2 can act through its receptor to enhance the *Clock*-*Bmal*-driven expression of its own gene. Should this occur within the SCN, it would generate an auto-amplification loop that might explain the massive daytime peak of prokineticin 2 expression, broadcasting an unequivocal midday chime across the hypothalamus and beyond. The mechanism of auto-amplification is unknown, but the activity of *Clock*-*Bmal* complexes can be regulated by physiological changes in redox potential⁷. So one possibility is that prokineticin amplifies its own production by modifying metabolism in the SCN.

How might these experimental findings from rodents² relate to people? In species that are active during the day, such as humans, the SCN clock mechanism cycles in phase with that of nocturnal rodents⁸: expression of *Period* is high during circadian day in both (but equates to wakefulness in diurnal species, compared with inactivity and sleep in nocturnal rodents). The genetic

sequences at which clock components such as Period regulate prokineticin 2 are conserved between humans and mice, so it is likely that prokineticin 2 levels will also be high in the daytime in humans. The expectation, therefore, is that by signalling daytime, SCN-derived prokineticin 2 would promote and consolidate human activity and wakefulness, rather than suppressing it — the opposite effect to that in rodents.

Beyond the SCN, circadian regulation of gastrointestinal function is well documented in people held in temporal isolation in experiments, and many of us are well aware of our own sanitary routines. Gastrointestinal dysfunction is an acute complaint of jet-lagged travellers, and a serious, chronic problem in shift workers. So it will be interesting to find out whether prokineticin signalling is involved in circadian regulation of the gut as

well as the brain, and indeed in other organ systems, such as the liver, that exhibit SCN-dependent circadian programmes of gene expression⁹. Cheng *et al.*'s results² offer enormous scope for intervening in sleep and other circadian disorders, and provide a new tool with which to explore the molecular and neurochemical control of complex behaviours. ■

Michael H. Hastings is in the Neurobiology Division, Laboratory of Molecular Biology, MRC Centre, Hills Road, Cambridge CB2 2QH, UK.
e-mail: mh105@cus.cam.ac.uk

1. Reppert, S. M. & Weaver, D. R. *Annu. Rev. Physiol.* **63**, 647–676 (2001).
2. Cheng, M. Y. *et al.* *Nature* **417**, 405–410 (2002).
3. Toh, K. L. *et al.* *Science* **291**, 1040–1043 (2001).
4. Mignot, E. *Neuropsychopharmacology* **25**, S5–S13 (2001).
5. Silver, R. *et al.* *Nature* **382**, 810–813 (1996).
6. Li, M. *et al.* *Mol. Pharmacol.* **59**, 692–698 (2001).
7. Rutter, J. *et al.* *Science* **293**, 510–514 (2001).
8. Mrosovsky, N. *et al.* *J. Biol. Rhythms* **16**, 471–478 (2001).
9. Akhtar, R. A. *et al.* *Curr. Biol.* **12**, 540–550 (2002).

Applied physics

Making sense of magnetic fields

Yeong-Ah Soh and Gabriel Aeppli

Although there have been dramatic advances in generating enormous magnetic fields, measuring these fields remains difficult. A non-metallic sensor based on a silver chalcogenide looks promising.

Over the past half-century, along with the growth in information technology using magnetic disk media, there has been a steady growth in the size of magnetic fields that can be routinely generated by wire coils. Today, steady fields approaching 20 Tesla (T) and pulsed fields as high as 40 T can be reached using relatively compact apparatus; specialized centres such as the US National High Magnetic Field Laboratory have routine access to fields as high as 60 T (for comparison, the Earth's magnetic field is 5×10^{-4} T). The growth in achievable magnetic-field strengths has created new frontiers for research, ranging from the investigation of novel quantum states to magnetic-resonance imaging.

With the increased range of magnetic field

available, there is a concomitant need for cheap and reliable magnetic-field sensors that work over a wide range of fields but that aren't sensitive to temperature. Over the past 15 years, the quest for low-field performance, driven largely by the multibillion-dollar magnetic data-storage industry, has produced a succession of devices based on magnetoresistance (MR) — the change in a material's resistance to electric-current flow caused by a magnetic field. Different devices exploit different types of magnetoresistance, such as giant (GMR)^{1–3}, colossal (CMR)^{4,5}, extraordinary (EMR)⁶ and tunnelling (TMR)⁷ magnetoresistance. GMR has, in fact, revolutionized the magnetic data-storage industry. On page 421 of this issue, Husmann *et al.*⁸ describe a new type of magnetoresistive

sensor. Based on a very different material, their device works over a wide range of magnetic-field strengths and temperatures, and its performance has little dependence on temperature.

The silver chalcogenide Ag₂Se forms the basis of the new device⁸. Figure 1 shows how this material performs when compared with other known magnetoresistive sensors. An earlier collaboration had accidentally discovered⁹ — during a search for voltage-induced flows of silver ions — that the resistivity of silver chalcogenides can be lowered to modest values by adding small amounts of excess silver to the compounds. At the same time, the compounds were found to have a nearly linear magnetoresistance. Husmann *et al.* show that this holds over a remarkably large range of fields (up to 50 T) and temperatures (1.5 K to 290 K), which makes the silver chalcogenides very attractive candidates for magnetic sensors.

The sensors can be very simple two-terminal devices: two conducting wires are attached to a block of silver chalcogenide, and the magnetic field is measured directly from the voltage drop between the wires as current flows between them. Existing devices called Hall sensors¹⁰, in which the field is proportional to the voltage induced transverse to a measuring current and to the magnetic field, are useful over a narrower temperature range than Ag₂Se and need to be carefully aligned with respect to the applied field. They also have high 'impedance', which slows down their response to rapidly varying fields, and restricts their use for pulsed fields.

Despite its applications, magnetoresistance is not completely understood. To describe the phenomenon, physicists start from the fundamental notions that electrical resistance arises from the scattering of electrons as they flow through a material and from the availability of quantum states for the electrons. A magnetic field can change the resistance in three ways: by altering the strength of the scattering; by modifying the trajectories followed by the electrons; or by changing the availability of quantum states for the electrons. The first two follow from the two key attributes of an electron, namely its spin and its charge, and the ability of magnetic fields to align spins and direct the motion of moving charges. The mechanisms of GMR and CMR are both based on spin-dependent scattering — spin alignment by an external field reduces the scattering probability of the electrons responsible for the current. The final mechanism, changing the availability of electron quantum states, is more subtle and can, in principle, be due to either spin (as it is in TMR) or charge motion.

In contrast to the mechanisms behind GMR, TMR and CMR, the magnetoresistance of the silver chalcogenides does not involve spin, because these materials are non-magnetic. It seems instead that the

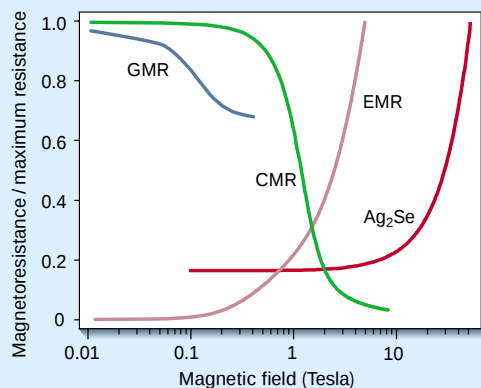


Figure 1 **Performance of magnetic-field sensors.** The dependence of magnetoresistance on applied magnetic field is shown for typical devices based on giant magnetoresistance (GMR) at a temperature of 295 K, colossal magnetoresistance (CMR) at 220 K and extraordinary magnetoresistance (EMR) at 300 K. Husmann *et al.*⁸ have devised a new sensor based on the silver chalcogenide Ag₂Se, which can be used to measure magnetic-field strengths as high as 50 Tesla. The data shown here were measured at a temperature of 290 K, but the device performs just as well over a wide temperature range, even down to just a few degrees above absolute zero. (Data derived from refs 3, 6, 8 and A. Biswas, personal communication.)

behaviour of the silver chalcogenides is analogous to that of the non-magnetic narrow-gap semiconductor indium antimonide (InSb), which displays EMR⁶ when metallic inhomogeneities are introduced. (In EMR, a magnetic field deflects the trajectories of moving electrons so that fewer pass through the metallic islands.) Inhomogeneity is a common factor between these two types of material. Stoichiometric silver chalcogenides do not show appreciable magnetoresistance⁹, and a marked increase in the magnetoresistance occurs only when the material is doped away from stoichiometry with excess silver.

Attempts to explain the unusual linear magnetoresistance in silver chalcogenides start with the idea that the excess silver atoms form metallic clusters and that the rest of the material has a very low concentration of mobile electrons¹¹. The metallic clusters are strikingly similar to the metallic inhomogeneity introduced by lithographic patterning that establishes the EMR effect in InSb. In both cases, the magnetoresistance is positive — the resistance increases with increasing magnetic-field strength. For inhomogeneities with small diameters, the magnetoresistance of InSb is linearly dependent on magnetic field up to the highest field value probed (5 T; ref. 6), in remarkable agreement with what is seen in the silver chalcogenides. But, as the diameter of the inhomogeneity increases, saturation of magnetoresistance becomes apparent at progressively smaller

field values. This suggests that the size and geometry of the metallic clusters may be crucial for the magnetoresistance of the silver chalcogenides.

The results reported by Husmann and colleagues⁸ are consistent with this interpretation, but other aspects of the data indicate that quantum effects¹¹ also contribute. The way forward is to perform controlled experiments in which the silver-cluster sizes are varied chemically and monitored by electron microscopy. These would shed light on whether the mechanism behind their magnetoresistance is indeed a natural, nanometre-scale version of what happens by brute-force engineering in EMR devices. ■

Yeong-Ah Soh and Gabriel Aeppli are at the NEC Research Institute, Princeton, New Jersey 08540, USA.

e-mails: soh@research.nj.nec.com; gabe@research.nj.nec.com

1. Grünberg, P., Schreiber, R., Pang, Y., Brodsky, M. B. & Sowers, H. *Phys. Rev. Lett.* **57**, 2442–2445 (1986).
2. Baibich, M. N. *et al. Phys. Rev. Lett.* **61**, 2472–2475 (1988).
3. Parkin, S. S. P., Li, Z. G. & Smith, D. J. *Appl. Phys. Lett.* **58**, 2710–2712 (1991).
4. von Helmolt, R., Wecker, J., Holzapfel, B., Schultz, L. & Samwer, K. *Phys. Rev. Lett.* **71**, 2331–2333 (1993).
5. Jin, S. *et al. Science* **264**, 413–415 (1994).
6. Solin, S. A., Thio, T., Hines, D. R. & Heremans, J. J. *Science* **289**, 1530–1532 (2000).
7. Julliere, M. *Phys. Lett. A* **54**, 225–226 (1975).
8. Husmann, A. *et al. Nature* **417**, 421–424 (2002).
9. Xu, R. *et al. Nature* **390**, 57–60 (1997).
10. Popović, R. S. *Hall Effect Devices: Magnetic Sensors and Characterization of Semiconductors* (Adam Hilger, Bristol, 1991).
11. Abrikosov, A. A. *Phys. Rev. B* **58**, 2788–2794 (1998).

Kidney function

Gateway to a long life?

Matthias A. Hediger

A molecule has been identified that seems to regulate the levels of uric acid in humans. The discovery has implications for understanding diseases such as gout and may even shed light on human longevity.

In vertebrates, uric acid is formed through the breakdown of unwanted purine molecules — the constituents of cellular energy stores such as ATP, and components of DNA and RNA. The breakdown process mainly occurs in the liver, muscles and intestine, and uric acid is then released into the bloodstream. Uric acid can be tremendously beneficial because it scavenges potentially harmful reactive oxygen species, inevitable by-products of glucose metabolism. Perhaps because of this, it also contributes to longevity in certain vertebrates. On the other hand, too much uric acid can also cause significant health problems, leading to kidney stones when it builds up in the kidneys and to gout when crystals accumulate in the joints. So the levels of uric acid in the blood must be tightly controlled. On page 447 of this issue, Enomoto and colleagues¹ describe how they identified

a long-hypothesized transporter molecule, found in the kidneys, that helps to maintain blood uric-acid levels in humans and might thereby maximize the benefits of uric acid.

Kidneys have two essential roles: they remove waste products from the blood, producing urine, and at the same time they regulate the blood levels of many important molecules, including salt and uric acid. Blood plasma is first filtered at the glomerulus, a network of capillaries, allowing water and small solutes, for instance uric acid — but not macromolecules such as proteins — to enter the many hollow tubules (nephrons) that make up the kidneys.

The fluid in the tubules forms urine and is excreted. But the cells that make up the tubule walls can reabsorb useful substances or allow further secretion of certain waste products. Reabsorbed substances pass

through the tubule cells back into the blood (Fig. 1, overleaf); waste products can be secreted by being transferred from blood directly into the tubule fluid. For example, uric acid is greatly reabsorbed in the upper (proximal) part of the tubules in humans, but is secreted into the tubule fluid in other species². The 'apical' membrane of tubule cells is that on the inside of the tubules. The 'basolateral' membrane is the blood-facing side. The reabsorption and secretion processes depend on specific transporter molecules that sit in these membranes.

Enomoto *et al.*¹ have now identified the transporter that reabsorbs uric acid. They started by scouring the draft human genome sequence for gene sequences related to those that encode known transporter proteins (specifically, those from the OAT protein family, called SLC22). The authors then found that one of these gene sequences, which they named URAT1, is expressed in human kidneys. Compellingly, the protein encoded by this gene is located in the apical membrane of proximal tubule cells¹. Moreover, frog eggs injected with URAT1-encoding RNA could transport uric acid.

So it seems likely that uric acid is transported across the apical membrane of human kidney tubule cells by URAT1 (with anions being transported back across the membrane, through URAT1, to maintain electrical balance). Uric acid is then moved across the basolateral membrane, into the blood, by an as-yet-unknown protein (Fig. 1). URAT1 is presumably absent in other mammals such as rabbits and pigs, as these species predominantly secrete uric acid². But in humans, uric-acid secretion is probably negligible, and URAT1 may be the major mechanism for regulating blood uric-acid levels.

The case put forward by Enomoto *et al.* for URAT1 as the apical uric-acid transporter is bolstered by their finding that it also interacts with several drugs and products of drug metabolism. For example, certain drugs that are used to treat inflammation or high blood pressure have undesirable side-effects on uric-acid excretion. 'Uricosuric' drugs block uric-acid reabsorption and 'anti-uricosuric' drugs enhance it. Enomoto *et al.* find that, in general, the uricosuric compounds directly inhibit URAT1 from the apical side whereas antiuricosuric drugs serve as the exchanging anion from inside tubule cells, thereby stimulating reabsorption.

The authors also studied patients with idiopathic renal hypouricaemia — a disorder that is primarily characterized by exercise-induced acute renal failure, triggered by the increased production of uric acid and reactive oxygen species that occurs in muscle during exercise³. An inability to reabsorb uric acid leads to the accumulation of uric-acid crystals in kidney tubules; this, together with exposure of the kidneys to reactive

Daedalus

Vegetarian meat

Last week Daedalus was enlarging birds' eggs by adding an expanded ceramic shell. DREADCO biologists are now extending this idea. They recall the old trick of making an egg plastic with vinegar, and re-setting it inside a bottle. This suggests that the nucleation of an egg shell depends on special proteins which can be interfered with. DREADCO sees fowl as rather primitive egg-formers, ripe for improvement. Hens will accept almost any source of calcium for egg shells, even calcium carbide, which is lethal to them. Inside the hen, the hard shell is put on the egg last. It can be coloured brown, by a special dye fed to the birds. Hens even waste their resources on eggs if these are not fertilized.

There is, of course, a commercial target: the 20% of the British population who are vegetarians. As an egg develops, its proteins change to meat, and DREADCO wants to sell meaty vegetarian eggs. A plastic shell would let the chick inside develop almost to adulthood still unconscious. It never has to wake up, emerge or grow feathers. It can breathe, for an egg is porous, even more so if it is plastic and as it stretches. The egg could be partly immersed in a nutrient broth, or maybe injected with one, so that a small initial chick could get bigger.

So the DREADCO team hope to make their hens lay stretchable, plastic eggs. They see the brittle, rigid, calcium carbonate egg shell as an evolutionary step backwards. Birds are descended from reptiles, whose eggs are leathery. Genes from closely related reptiles might make birds' eggs plastic again. Alternatively, simply feeding the birds on calcium salts of plastic acids, such as polyacrylic or polymethacrylic, might do the trick. And if the resulting eggs are incomplete or somehow unviable, like the usual unfertilized eggs, this doesn't matter if the hens can be bred some other way.

Eggs as currently sold are very remote from life. The DREADCO team hope to produce a bigger but equally artificial plastic egg. The ideal product would contain not a bird, but the meat elements from which one might be constructed. The most committed vegetarian should accept in a plastic envelope meat that had never been conscious. But for how long should the egg be grown? Should a nutrient broth be injected? Ideally, the consumer will open the egg, dissect out the meaty, edible section and reject the rest, without even thinking of the bird that might have been: any more than we think this of an egg.

David Jones

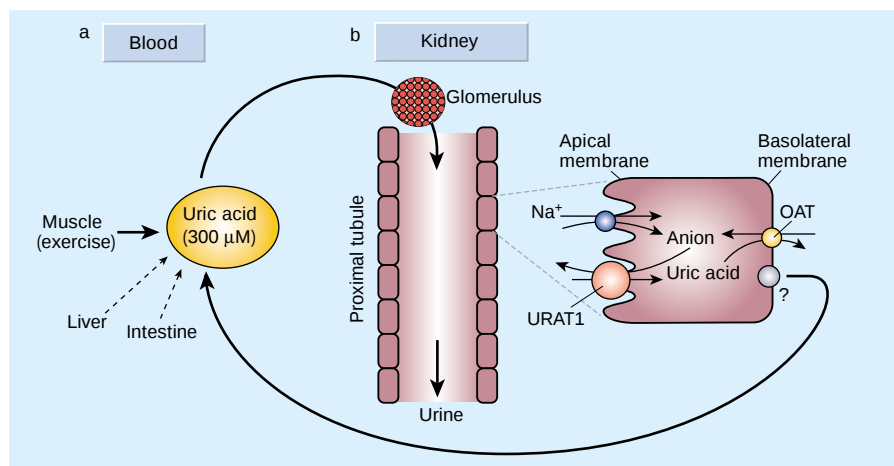


Figure 1 Pathways for uric-acid transport. a, Uric acid is produced as the major end-product of purine metabolism by muscles, liver and intestine. b, After the blood is filtered at the glomerulus, the resulting fluid enters the tubules of the kidneys to allow waste substances to be excreted as urine. Many substances, however, such as uric acid, are useful and so are reabsorbed. This is achieved by the cells that line the walls of the tubules, as seen in detail at the right. As shown by Enomoto *et al.*¹, the transporter URAT1 mediates uptake of uric acid across the apical membrane, in exchange for anions, which enter cells either by Na⁺-coupled transport through the apical membrane or organic anion transport (through OAT proteins) across the basolateral membrane. Uric acid leaves the cell across the basolateral membrane through an unknown channel or transporter, possibly in exchange for anions through an OAT protein.

oxygen species, causes death of tubule cells. Without exercise, these patients can live normally, except for an increased occurrence of kidney stones. Enomoto *et al.* find that these patients have defects in the URAT1 gene — yet more evidence that this indeed encodes the uric-acid transporter.

These results¹ have several implications. For example, natural person-to-person variations in the URAT1 gene might contribute to certain forms of gout, if such variations lead to greater reabsorptive activity. The prevalence of gout is known to increase with blood uric-acid concentration, along with several other factors including age, body weight, blood pressure, alcohol intake and consumption of purine-rich foods such as meat⁴. There is also evidence that higher blood levels of uric acid are associated with a greater risk of heart disease (although it is unclear whether uric acid contributes directly⁵), so natural variations in URAT1 might be important here, too.

Finally, URAT1 might contribute to human longevity. An important factor in the lengthening of primate lifespans has been the evolution of mechanisms that protect against reactive oxygen species. Such mechanisms include antioxidant compounds such as vitamins A, C and E^{6,7}. Uric acid is thought to be another antioxidant, as suggested by the fact that birds are very long-lived for their body size — despite high metabolic rates, body temperatures and blood glucose levels⁸ — and also produce uric acid as the major end-product of the metabolism of protein-derived nitrogen.

In contrast to birds, mammals produce urea rather than uric acid as the end-product

of such metabolism. But uric acid is still produced, from purines, and notably it accumulates in humans to about ten times the levels in most other mammals². Blood uric-acid levels in humans are also about six times those of vitamin C (about 300 μM compared with some 50 μM)⁹. So uric acid is thought to be a primary antioxidant in human blood. The evolution of URAT1, and the loss of the ability to secrete uric acid and of the liver uric-acid-degrading enzyme, was presumably important in producing these levels. It will be interesting to see whether variations or mutations in URAT1 correlate with variations in human lifespans.

Uric acid clearly has many beneficial effects. But, in combination with genetic or environmental factors, it can also be detrimental when it accumulates at high levels. So the human uric-acid transporter URAT1 may prove an interesting target for future drug development.

Matthias A. Hediger is in the Membrane Biology Program at the Brigham and Women's Hospital and the Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115, USA.

e-mail: mhediger@rics.bwh.harvard.edu

- Enomoto, A. *et al.* *Nature* **417**, 447–452 (2002).
- Roch-Ramel, F. & Guisan, B. *News Physiol. Sci.* **14**, 80–84 (1999).
- Kikuchi, Y. *Clin. Nephrol.* **53**, 467–472 (2002).
- Wortmann, R. L. & Kelley, W. N. in *Ruddy's Textbook of Rheumatology* 6th edn 1339–1389 (W. B. Saunders, Philadelphia, 2001).
- Johnson, R. J. *et al.* *Am. J. Kidney Dis.* **33**, 225–234 (1999).
- Ames, B. N., Cathcart, R., Schwiers, E. & Hochstein, P. *Proc. Natl Acad. Sci. USA* **78**, 6858–6862 (1981).
- Mecocci, P. *et al.* *Free Rad. Biol. Med.* **28**, 1243–1248 (2000).
- Simoyi, M. F., Van Dyke, K. & Klandorf, H. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **282**, R791–R796 (2002).
- Hediger, M. A. *Nature Med.* **8**, 445–446 (2002).

Obituary

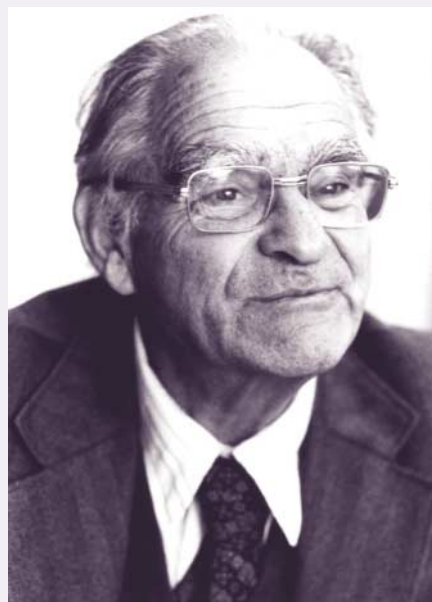
Victor F. Weisskopf (1908–2002)

Victor Weisskopf, who died on 21 April at the ripe age of 93, had, as he liked to put it, “lived a happy life in a dreadful century”. He knew what he was talking about. He had been a major player in a great voyage of discovery, and was always at the centre of a loving family and a vast circle of good friends. But he had also seen totalitarianism in Germany and Russia, and witnessed the first man-made nuclear explosion.

In 1928, Weisskopf left his native Vienna to study physics at Göttingen, Germany, where quantum mechanics had been born just three years before. It was already clear that the new mechanics could describe, at least in principle, any phenomenon involving particles moving with velocities that are small compared with the speed of light. Now the question was whether a successful marriage of quantum mechanics with relativity could be arranged. Paul Dirac had responded by inventing quantum electrodynamics (QED), but it was soon realized that this theory faced potentially fatal problems. It was to this fundamental issue that Weisskopf repeatedly devoted himself. Today, QED and the theory of gravitation are the most accurately tested theories of physics — an outcome due in good part to Weisskopf’s contributions. Willis Lamb, who did the most important experiment in the history of the subject, wrote that he based his analysis on “the 1927–34 formulation of QED due to Dirac, Heisenberg, Pauli and Weisskopf”.

Weisskopf’s first landmark result concerned the spectral lines resulting from the radiation emitted by electrons in the atom. The intrinsic width of spectral lines is due to the interaction of the atomic electrons with the radiation they emit. Working with Eugene Wigner, and at the age of just 22, Weisskopf devised the quantum theory of this effect.

The energy of a charge under the influence of its own field posed a much deeper problem. In 1934, with some help from Wendell Furry, Weisskopf discovered that in QED this ‘self-energy’ is far better behaved as the diameter of the charge shrinks to zero than it is in classical physics. This behaviour is central to today’s highly successful ‘standard model’ of elementary particles. Another vital contribution by the young Weisskopf was the demonstration, with Wolfgang Pauli, that relativity and quantum mechanics do not require charged particles to have spin — until then, this



Quantum theorist and inspiring leader

had been a misconception widely inferred from Dirac’s theory.

After Göttingen, Weisskopf had perhaps the most spectacular postdoctoral career of modern physics, holding appointments with Werner Heisenberg, Erwin Schrödinger, Pauli and Niels Bohr. In 1937, he left Europe for a position at the University of Rochester and began a long-term involvement with nuclear physics. This culminated in the book *Theoretical Nuclear Physics*, written with John Blatt and published in 1952, which for decades was the undisputed bible of the subject. But what had begun as a harmless intellectual pursuit assumed a very different tone when he joined the atomic bomb project at Los Alamos. There he became deputy to Hans Bethe, head of the theoretical-physics division.

At the war’s end he joined the faculty at Massachusetts Institute of Technology. Here, students and postdocs were exposed to Weisskopf’s exceptional talent for the intuitive argument, and also enjoyed the warmth of the atmosphere he created. New experimental techniques had emerged from the wartime radar project, and among the first consequences was Lamb’s discovery that the hydrogen spectrum was in stark conflict with existing theory. Weisskopf and his student Bruce French made the first complete calculation, but it disagreed with the slightly later result found by the young

geniuses Richard Feynman and Julian Schwinger with their powerful, separate reformulations of QED. Weisskopf, who suffered from a lack of confidence in mathematics, did not publish, assuming that he and French were wrong. They were not — the other two had made the same subtle mistake.

The deep wounds inflicted on European science by the Nazis and the nuclear arms race motivated many of Weisskopf’s thoughts and actions after the detonation of the atom bomb over Hiroshima. In 1946 he joined a small committee chaired by Einstein, the first of many platforms from which he spoke of the danger posed by nuclear weapons. He was a co-founder of the Federation of American Scientists and the Union of Concerned Scientists. Starting with the first Pugwash conference in 1956, he pioneered relations with Soviet scientists in pursuit of arms control. And it is no coincidence that Pope John Paul II began to make forceful and widely publicized statements against the nuclear arms race soon after Weisskopf’s election to the Pontifical Academy.

Weisskopf never severed his strong ties with Europe. He was delighted to be named director-general of CERN (the European laboratory for particle physics in Geneva, Switzerland) in 1961, having answered questions about his administrative experience with “none whatever, my greatest strength”. He proved to be an inspiring and effective leader, not only of CERN but of high-energy physics worldwide. Before long, CERN and European high-energy physics were on a par with competing research in the United States, in good part thanks to his controversial and ground-breaking decision to build the world’s first proton–proton collider.

A sketch of ‘Viki’ (as he was known everywhere) as a scientist does not begin to give a picture of the man. He radiated emotional warmth combined with intellectual discipline and breadth — a rigorous romantic. He was an accomplished pianist and loved music as much as physics, playing excerpts from Haydn’s *Creation* when giving popular lectures on the early Universe. He called his popular exposition of science *Knowledge and Wonder*. Viki exemplified the vitality and imagination that produced one of history’s great intellectual revolutions, and gave it a human face.

Kurt Gottfried

Kurt Gottfried is in the Laboratory of Nuclear Studies, Cornell University, Ithaca, New York 14853, USA.

e-mail: kg13@cornell.edu

A Möbius strip of single crystals

A crystalline ribbon of niobium and selenium can be coaxed into a novel topology.

A Möbius strip is produced by twisting a ribbon of material through 180° and joining its two ends, resulting in a distinct, one-sided topology^{1–4}. Here we describe a Möbius structure formed by crystals of a compound of niobium and selenium, NbSe_3 . It is surprising that a crystalline ribbon should adopt this exotic topology in view of its inherent rigidity, which would be expected to prevent it from either bending or twisting.

NbSe_3 is a typical low-dimensional inorganic conductor, which shows charge-density-wave (CDW) phase transitions at 58 K and 145 K. Its crystals are synthesized by chemical-vapour transport and assume fibrocristalline ribbon and whisker shapes. We have been able to create a Möbius strip of single NbSe_3 crystals by modifying the conventional growth conditions.

We soaked a mixture of Se and Nb powder at 740°C in an evacuated quartz tube for a period varying from a few hours to a few days. The growth conditions differed from those that are typically used, in that they relied on a furnace that had a large temperature gradient; this creates a crucial non-equilibrium state inside the quartz tube in which selenium circulates through vapour, mist and liquid (droplet) phases.

Figure 1 shows scanning electron microscopic (SEM) images of different topologies that are assumed by NbSe_3 crystals produced under these conditions. Ring diameter and width are typically $100\ \mu\text{m}$ and $1\ \mu\text{m}$, respectively (Fig. 1a); Möbius crystals are typically around $50\ \mu\text{m}$ in diameter and less than $1\ \mu\text{m}$ in width (Fig. 1b); and ‘figure-of-eight’ crystal strips (Fig. 1c), with a double twist, have a circumference of about $200\ \mu\text{m}$ and a width of $1\ \mu\text{m}$. We confirmed by X-ray diffraction and transmission electron microscope diffraction that the ring and Möbius structures are equivalent to a conventional crystalline ribbon with single-phase and monoclinic properties.

The mechanism of ring formation by the ribbon-shaped NbSe_3 crystals depends on their being bent by the surface tension of the viscous selenium droplet on which they have grown. The ribbon grows along the equator of the droplet to minimize bending energy, finally meeting its own tail to form a perfectly seamless ring (Fig. 1d).

Growth of the Möbius strip is difficult by comparison because of the twisting involved. Bending of a bar or beam is accompanied by twisting, and the low symmetry in a monoclinic crystal also promotes bending with twisting⁵, as shown

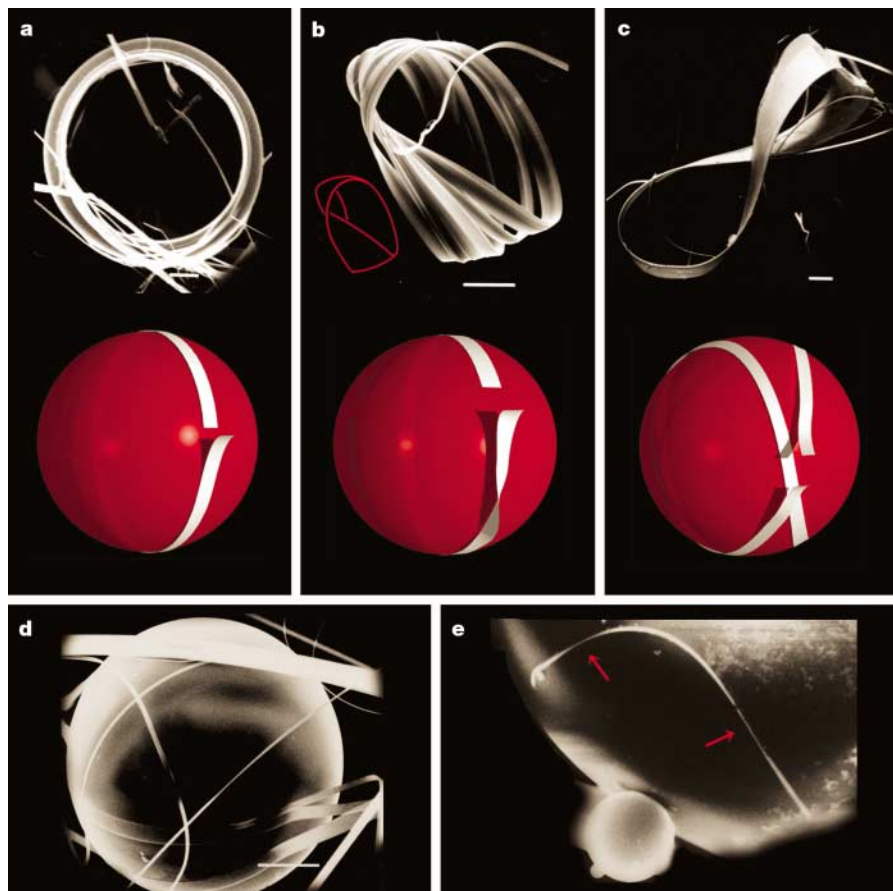


Figure 1 Scanning electron microscopic images of NbSe_3 crystal topology. **a–c**, The three types of topology, classified by the nature of their twisting (shown schematically beneath images; white ribbons represent NbSe_3 crystals, red spheres represent selenium droplets). **a**, Ring structure (0π twist). **b**, Möbius strip (1π twist). **c**, Figure-of-eight strip (2π twist). **d**, NbSe_3 fibres (white streaks) bend into rings around a selenium drop of diameter about $50\ \mu\text{m}$. The NbSe_3 ribbon is spooled onto a selenium droplet by surface tension until its ends bind together. **e**, High-magnification image showing a twist in a ribbon of NbSe_3 crystals; spooling of the ribbon can produce a twist, as in the formation of a Möbius crystal, owing to its anisotropic elastic properties. Scale bars, $10\ \mu\text{m}$.

in Fig. 1e. Droplet rotation, which we observed frequently, could also help to produce the twist.

The figure-of-eight crystal strips arise as a result of either double encircling or double twisting. According to White's theorem⁶, a doubly encircling loop is topologically equivalent to a 2π -twisted loop, which is analogous to a ring conformer of DNA⁷.

These different topological arrangements of NbSe_3 all show the characteristic CDW phase transitions, indicating that the samples are crystalline and relatively ordered, and that the CDW displacements are identical to those in conventional, untwisted crystals^{8,9}. This was demonstrated by satellites in the electron-diffraction patterns (CDW wave vector $\mathbf{Q}_1 = 0.24 \pm 0.01$), by anomalies due to CDW phase transitions in the temperature dependence of the resistivity at 145 K and 54 K, and by non-linear conduction due to CDW sliding.

We have developed a new growth technique for topological crystals by using a spherical droplet as a spool. This technique could be extended to a wide range of materials, and we have already obtained topological variants of tantalum compounds with selenium and sulphur (TaSe_3 and TaS_3). It may also be possible to grow crystals to a desired size by controlling the non-equilibrium conditions inside the furnace and thus the size of the droplets. Our crystal forms offer a new route to exploring topological effects in quantum mechanics^{10–13}, as well as to the construction of new devices.

Satoshi Tanda*, **Taku Tsuneta***,
Yoshitoshi Okajima*, **Katsuhiko Inagaki***,
Kazuhiko Yamaya*, **Noriyuki Hatakenaka†**

*Department of Applied Physics, Hokkaido University, Sapporo, Hokkaido 060-8628, Japan
e-mail: tanda@eng.hokudai.ac.jp

†NTT Basic Research Laboratories, Atsugi, Kanagawa 243-0198, Japan

- Walba, D. M., Richards, R. M. & Haltiwanger, R. C. *J. Am. Chem. Soc.* **104**, 3219–3221 (1982).
- Mila, F., Stafford, C. & Capponi, S. *Phys. Rev. B* **57**, 1457–1460 (1998).
- Hayashi, M. & Ebisawa, H. *J. Phys. Soc. Jpn* **70**, 3495–3498 (2001).
- Kaneda, K. & Okabe, Y. *Phys. Rev. Lett.* **86**, 2134–2137 (2001).
- Hearmon, R. F. S. *An Introduction to Applied Anisotropic Elasticity* (Oxford Univ. Press, London, 1961).
- White, J. H. *Am. J. Math.* **91**, 693–728 (1969).
- Vologodskii, A. V., Anshelevich, V. V., Lukashin, A. V. &

- Frank-Kamenetskii, M. D. *Nature* **280**, 294–298 (1979).
- Tsutsui, K. et al. *Phys. Rev. Lett* **39**, 1675–1676 (1977).
- Monceau, P. *Electronic Properties of Inorganic Quasi-one-dimensional Compounds* (Reidel, Dordrecht, 1985).
- Berry, M. V. *Proc. R. Soc. Lond. A* **392**, 45–57 (1984).
- Ando, T., Nakanishi, T. & Saito, R. *J. Phys. Soc. Jpn* **67**, 2857–2862 (1998).
- Iijima, S. *Nature* **354**, 56–58 (1991).
- Hyde, S. T. *Phys. Chem. Miner.* **20**, 190–200 (1993).

Competing financial interests: declared none.

Endosymbiotic bacteria

GroEL buffers against deleterious mutations

GroEL, a heat-shock protein that acts as a molecular chaperone¹, is overproduced in endosymbiotic but not in free-living bacteria^{2–4}, presumably to assist in the folding of conformationally damaged proteins. Here we show that the overproduction of GroEL in *Escherichia coli* masks the effects of harmful mutations that have accumulated during a simulated process of vertical transmission. This molecular mechanism, which may be an adaptation to the bacterium's intracellular lifestyle, is able to rescue lineages from a progressive fitness decline resulting from the fixation of deleterious mutations under strong genetic drift^{5,6}.

Endosymbiotic bacteria have small population sizes, do not undergo recombination, and are maternally transmitted through tight population bottlenecks⁷, causing the fixation of deleterious mutations due to genetic drift and hence an irreversible decline in fitness⁸. However, endosymbiosis is surprisingly stable and persists over long periods⁹, which has led to the suggestion that *groE* (the GroEL-encoding operon) could be buffering the mutational loss of functionally active proteins because, unlike other endosymbiont genes, it is subject to strong purifying selection⁵.

To test this idea, we investigated the effects of overexpression of *groE* in a set of *E. coli* strains (a free-living bacterium close to several endosymbionts⁹) with mutations randomly accumulated throughout the genome. These spontaneous mutations were fixed by random genetic drift in a process that simulated the vertical transmission of a single endosymbiont between hosts.

We studied the accumulation of mutations in 12 replicate lines of two *E. coli* B genotypes, one of which had a 3.3-fold-increased mutation rate¹⁰. These genotypes had already adapted to a simple environment (DM25 medium) for 10,000 generations¹¹, suggesting that mutation accumulation might result in a decline in fitness. After 3,240 generations of mutation accumulation, we measured the fitness of the strains evolved on DM25 relative to their respective ancestors. As expected, the

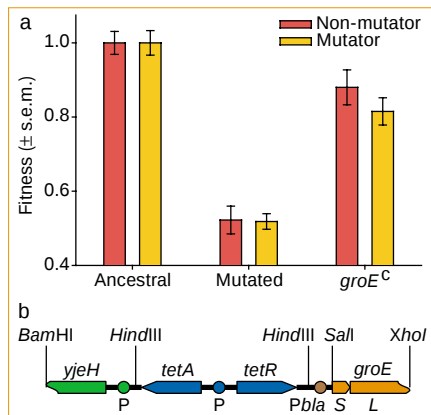


Figure 1 Effect of the overexpression of the *groE* operon on the fitness of randomly mutated strains of *Escherichia coli*. **a**, The fitness values for the ancestral strains of mutators and non-mutators are the same (left); 'mutated', fitness loss after 3,240 generations of mutation accumulation; 'groE⁺', fitness recovery of strains in which *groE* is overexpressed. Culture conditions, fitness assays and phenotypic markers are described elsewhere¹¹, as is the mutation-accumulation protocol¹². **b**, The regulated *E. coli* *groE* operon was changed to a constitutive form (*groE*⁺) by putting the structural genes under the control of the β -lactamase gene promoter (*Pbla*). Recombinant genotypes were obtained by transduction using P1vir (ref. 13). The *groE*⁺ construct contains part of the flanking *yjeH* gene, the inducible tetracycline-resistance operon (*tetA/R*) for transductant selection, the *Pbla* promoter, and the *groE* coding regions (S and L). Restriction-enzyme cutting sites are indicated; P, promoter. Further details are available from the authors.

mutated strains lost fitness (Fig. 1a) as a result of the accumulation of deleterious mutations ($\bar{W}=0.5168 \pm 0.0159$ for mutators; $\bar{W}=0.5208 \pm 0.0087$ for non-mutators; paired *t*-tests, $P < 0.0001$).

We then replaced the regulated *groE* operon of the mutated strains with a constitutive allele (*groE*⁺; Fig. 1b). To estimate the fitness of the *groE*⁺ strains, we ran competition experiments against the corresponding ancestors on DM25. Surprisingly, none of the *groE*⁺ strains reached the expected cell density during growth overnight. A simple explanation could be that an overproduction of GroEL of about 86 ± 16 -fold is deleterious because it diverts amino acids away from other cellular functions.

To test this, we grew each *groE*⁺ strain and its ancestor in DM25 supplemented with increasing concentrations of tryptone (a mixture of peptides and amino acids). We found that *groE* overexpression was deleteri-

ous in the presence of small amounts of tryptone, but that there were no significant differences in cell density between *groE*⁺ strains and their non-mutated ancestors at tryptone concentrations of 0.1% or higher. In endosymbionts, amino acids are not limiting because these are abundant in their intracellular environment⁹.

To determine whether fitness estimates depend on the environment, we estimated fitness on DM25 and on DM25 + 0.1% tryptone. Both fitness estimates were correlated (partial correlation test, $P < 0.0001$).

Figure 1a shows that the average fitness of the *groE*⁺ strains derived from non-mutator strains ($\bar{W}=0.8801 \pm 0.0214$) was 75.9% greater than that of the mutated strains (paired *t*-test, $P < 0.0001$), but was 12% less than that of the ancestors ($P = 0.0002$); the average fitness of the *groE*⁺ strains derived from evolved mutators was $\bar{W}=0.8152 \pm 0.0167$, which is 61.6% greater than that of the mutated strains (paired *t*-test, $P < 0.0001$) but 18.48% less than that of the ancestral strains ($P < 0.0001$).

Is fitness recovery a result of the buffering of deleterious effects by GroEL, or is it simply a general benefit associated with increased concentrations of GroEL? In favour of the first possibility, the advantage of GroEL overproduction is evident only in an amino-acid-rich environment; also, if mutation compensation is occurring, we would expect a positive correlation between the extents of fitness loss and recovery, as evidenced by their partial correlation (1-tailed test, $P = 0.0089$). We conclude that GroEL overexpression is likely to be of help in maintaining these endosymbionts by protecting them against the harmful effects of accumulated mutations.

Mario A. Fares, Mario X. Ruiz-González, Andrés Moya, Santiago F. Elena, Eladio Barrio

Institut Cavanilles de Biodiversitat i Biologia Evolutiva i Departament de Genètica, Universitat de València, PO Box 22085, 46071 València, Spain
e-mail: eladio.barrio@uv.es

- Sigler, P. B. et al. *Annu. Rev. Biochem.* **67**, 581–608 (1998).
- Sato, S. & Ishikawa, H. *J. Bacteriol.* **179**, 2300–2304 (1997).
- Aksoy, S. *Insect Mol. Biol.* **4**, 32–29 (1995).
- Charles, H., Heddi, A., Guillaud, J., Nardon, C. & Nardon, P. *Biochem. Biophys. Res. Commun.* **239**, 769–774 (1997).
- Moran, N. A. *Proc. Natl Acad. Sci. USA* **93**, 2873–2878 (1996).
- Rutherford, S. L. & Lindquist, S. *Nature* **396**, 336–342 (1998).
- Funk, D. J., Wernegreen, J. J. & Moran, N. A. *Genetics* **157**, 477–489 (2001).
- Muller, H. J. *Mutat. Res.* **1**, 2–9 (1964).
- Moran, N. A. & Wernegreen, J. J. *Trends Ecol. Evol.* **15**, 321–326 (2000).
- de Visser, J. A. G. M., Zeyl, C., Gerrish, P. J., Blanchard, J. L. & Lenski, R. E. *Science* **283**, 404–406 (1999).
- Lenski, R. E., Rose, M. R., Simpson, S. C. & Tadler, S. C. *Am. Nat.* **138**, 1315–1341 (1991).
- Kibota, T. T. & Lynch, M. *Nature* **381**, 694–696 (1996).
- Zyskind, J. W. & Bernstein, S. I. *Recombinant DNA Laboratory Manual* (Academic, San Diego, 1989).

Competing financial interests: declared none.

Comparative assessment of large-scale data sets of protein–protein interactions

Christian von Mering*, Roland Krause†, Berend Snel*, Michael Cornell‡, Stephen G. Oliver‡, Stanley Fields§ & Peer Bork*

* European Molecular Biology Laboratory, Meyerhofstrasse 1, 69012 Heidelberg, Germany

† Cellzome AG, Meyerhofstrasse 1, 69117 Heidelberg, Germany

‡ School of Biological Sciences, University of Manchester, 2.205 Stopford Building, Oxford Road, Manchester M13 9PT, UK

§ Departments of Genome Sciences and Medicine, Howard Hughes Medical Institute, University of Washington, Box 357730, Seattle, Washington 98195, USA

Comprehensive protein–protein interaction maps promise to reveal many aspects of the complex regulatory network underlying cellular function. Recently, large-scale approaches have predicted many new protein interactions in yeast. To measure their accuracy and potential as well as to identify biases, strengths and weaknesses, we compare the methods with each other and with a reference set of previously reported protein interactions.

For an increasing number of organisms, we can now list the genes and encoded proteins^{1,2}. It is the proteins that execute the genetic programme, and those that are actually produced by a cell at any given time constitute its ‘proteome’³. The proteome is much more dynamic than the genome: it changes during development and in response to external stimuli, and the proteins form large interaction networks, in which they regulate and support each other^{4–6}. To fully understand the cellular machinery, simply listing the proteins is not enough—all the interactions between them need to be delineated as well.

Traditionally, protein interactions have been studied individually by genetic, biochemical and biophysical techniques. However, the speed with which new proteins are being discovered or predicted has created a need for high-throughput interaction-detection methods. Consequently, in the last two years, methods have been introduced that can globally tackle the problem, resulting in a vast amount of interaction data^{7–14}. To use these data efficiently, a critical evaluation of their accuracy, biases, overlaps and complementarities is essential. Here, we analyse data sets from yeast two-hybrid systems^{7,8}, protein complex purification techniques using mass spectrometry^{10,11}, correlated messenger RNA expression profiles^{15,16} and genetic interaction data^{9,17}, as well as ‘*in silico*’ (computed) interaction predictions derived from gene context analysis (gene fusion^{18,19}, gene neighbourhood^{14,20} and gene co-occurrences or phylogenetic profiles^{21,22}). We compare the methods with each other, focusing on the yeast proteome (for details of the methods, see Box 1). Although all of these techniques can be used for interaction prediction, their goals are different. Yeast two-hybrid and mass spectrometry techniques aim to detect physical binding between proteins, whereas genetic interactions, mRNA coexpression and *in silico* methods seek to predict functional associations, for example, between a transcriptional regulator and the pathway it controls. In many cases, however, such functional associations do take the form of physical binding^{20,23}.

Comparing interaction data is difficult, because they are often derived under different conditions, come in different formats (see Box 2), and need to be benchmarked against a trusted reference set. For this study, we chose binary interactions as the common unit of analysis, and we relied on manually curated catalogues of known protein complexes (Munich Information Center for Protein Sequences, MIPS¹⁷, and the Yeast Proteome Database, YPD²⁴) as the trusted reference.

Overlaps and complementarities

About 80,000 interactions between yeast proteins are currently available from the different high-throughput methods (Fig. 1) (the exact number depends on filtering criteria). Of these, only a surprisingly small number (~2,400) is supported by more than one method. There are three possible explanations for this: the methods may not have reached saturation; many of the methods may produce a significant fraction of false positives; and some methods may have difficulties for certain types of interactions, resulting in complementarities between the methods.

We note, for example, that each technique produces a unique distribution of interactions with respect to functional categories of interacting proteins (Fig. 1). These differences in coverage suggest that the methods have specific strengths and weaknesses. The data sets based on purified complexes, for example, predict relatively few interactions for proteins involved in transport and sensing (possibly because these are enriched in transmembrane proteins, which are more difficult to purify). Similarly, interactions detected by the yeast two-hybrid technology largely fail to cover certain categories; for example, proteins involved in translation are found comparatively less often than by other methods.

As a specific example for the complementarity between the data sets, we considered glycine decarboxylase, which is a well-characterized multi-enzyme complex needed when glycine is used as a one-carbon source^{25–27}. It consists of the four proteins Gcv1, Gcv2, Gcv3 and Lpd1. This complex is not detected in the systematic purification of complexes¹⁰ (using tandem affinity purification (TAP)-tagged Gcv3 as a bait), presumably because it is not expressed in yeast cells grown on rich medium. Three of the four proteins, however, can be confidently linked to each other through the remaining data sets, by a total of nine links involving synexpression (correlated mRNA expression) and *in silico* predictions (see Supplementary Information).

A converse example is the protein PPH3, for which neither *in silico* methods nor synexpression studies predict any interaction, but which is nevertheless consistently detected in a protein complex by mass spectrometry. PPH3 is a protein phosphatase distantly related to PP2A; very little is known about its function or interaction partners^{28,29}. In the high-throughput data, two uncharacterized proteins (YBL046W and YNL201C) are invariably found with PPH3 in four independent mass-spectrometry purifications, thereby clearly defining a protein complex (the two are also joined by a two-hybrid interaction).

Even data sets based on the same technique can complement each other to some extent. The two mass-spectrometry approaches, for

example, differ in how they express the tagged 'baits': the TAP approach¹⁰ relies on genomic integration (using the endogenous promoter), whereas the approach of high-throughput mass-spectrometry protein complex identification (HMS-PCI)¹¹ employs

Box 1 High-throughput methods for detecting protein interactions

Yeast two-hybrid assay. Pairs of proteins to be tested for interaction are expressed as fusion proteins ('hybrids') in yeast: one protein is fused to a DNA-binding domain, the other to a transcriptional activator domain. Any interaction between them is detected by the formation of a functional transcription factor^{7,8,41,42}. Benefits: it is an *in vivo* technique; transient and unstable interactions can be detected; it is independent of endogenous protein expression; and it has fine resolution, enabling interaction mapping within proteins. Drawbacks: only two proteins are tested at a time (no cooperative binding); it takes place in the nucleus, so many proteins are not in their native compartment; and it predicts possible interactions, but is unrelated to the physiological setting.

Mass spectrometry of purified complexes. Individual proteins are tagged and used as 'hooks' to biochemically purify whole protein complexes. These are then separated and their components identified by mass spectrometry. Two protocols exist: tandem affinity purification (TAP)^{10,43}, and high-throughput mass-spectrometric protein complex identification (HMS-PCI)^{11,44}. Benefits: several members of a complex can be tagged, giving an internal check for consistency; and it detects real complexes in physiological settings. Drawbacks: it might miss some complexes that are not present under the given conditions; tagging may disturb complex formation; and loosely associated components may be washed off during purification.

Correlated mRNA expression (synexpression). mRNA levels are systematically measured under a variety of different cellular conditions, and genes are grouped if they show a similar transcriptional response to these conditions. These groups are enriched in genes encoding physically interacting proteins²³. Benefits: it is an *in vivo* technique, albeit an indirect one; and it has much broader coverage of cellular conditions than other methods. Drawbacks: it is a powerful method for discriminating cell states or disease outcomes, but is a relatively inaccurate predictor of direct physical interaction; and it is very sensitive to parameter choices and clustering methods during analysis.

Genetic interactions (synthetic lethality). Two nonessential genes that cause lethality when mutated at the same time form a synthetic lethal interaction. Such genes are often functionally associated and their encoded proteins may also interact physically. This type of genetic interaction is currently being studied in an all-versus-all approach in yeast⁹. Benefits: it is an *in vivo* technique, albeit an indirect one; and it is amenable to unbiased genome-wide screens.

In silico predictions through genome analysis. Whole genomes can be screened for three types of interaction evidence: (1) in prokaryotic genomes, interacting proteins are often encoded by conserved operons^{13,14,20}; (2) interacting proteins have a tendency to be either present or absent together from fully sequenced genomes^{21,22}, that is, to have a similar 'phylogenetic profile'; and (3) seemingly unrelated proteins are sometimes found fused into one polypeptide chain. This is an indication for a physical interaction^{18,19}. Benefits: fast and inexpensive *in silico* techniques; and coverage expands as more genomes are sequenced. Drawbacks: it requires a framework for assigning orthology between proteins, failing where orthology relationships are not clear; and so far it has focused mainly on prokaryotes.

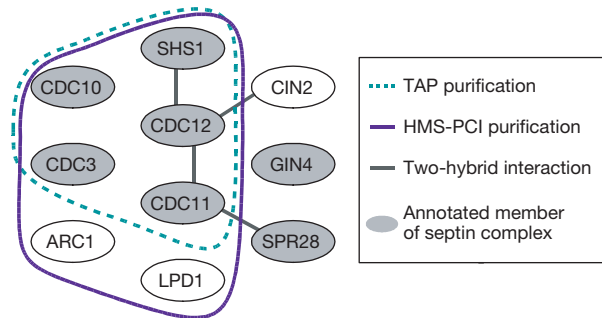
inducible overexpression. Both systems have their advantages: the endogenous promoter will often maintain the stoichiometry of interacting proteins and minimize artificial overexpression, whereas inducible expression allows the purification of proteins that are normally not expressed under laboratory conditions, and can be used in species where genomic integration is not feasible, such as the human. Of the proteins that could not be expressed in the TAP approach, 90 were successfully purified by HMS-PCI, producing more than 600 interactions.

Benchmarking high-throughput interactions

When assessing the quality of interaction data, coverage and accuracy need to be considered together. A data set of high coverage is not very useful if its accuracy is low (that is, it contains many false positives), and vice versa. Comparing the data with a reference set of trusted interactions allows the estimation of lower limits for accuracy and coverage. Figure 2 summarizes how these values relate to each other for the different data sets. Such a comparison can provide only a very rough picture, because it is a snapshot of ongoing efforts, because it is based on a particular framework for

Box 2 Counting interactions

Interaction data come in two formats: binary interactions or groups of interacting partners, shown here for a known protein complex, and the high-throughput data that support it (the septin complex, not all interactions are indicated).



For any quantitative comparison of data sets, a common unit of analysis needs to be defined, which in this case means focusing either on proteins or on binary interactions. It makes a difference: for the example shown, the two-hybrid data have a false-positive rate of 1 in 5 when counting proteins, but it is 1 in 4 when counting interactions. Counting proteins seems more intuitive at first glance, but it becomes complicated when comparing more than two partially overlapping experiments. Also, counting proteins has a lower resolution because it does not discriminate whether a protein is involved in one or more interactions.

When focusing on interactions, data that come in groups need to be expanded to all possible binary interactions within a group, and can then be compared with other data sets. For example, consider the overlap of the two mass-spectrometry approaches^{10,11}. When counting interactions, the overlap stands at 1,728 shared binary interactions (which is 27.5% of the TAP interactions¹⁰ and 19.2% of the HMS-PCI interactions¹¹, considering only proteins present in both data sets). This overlap is more difficult to define when counting proteins, and can be as high as 42% of the TAP data and 33% of the HMS-PCI data. (This depends on how one counts (see Supplementary Information). The largest overlap between two purifications is 18 proteins.) Importantly, the overall trend stays the same, irrespective of the unit of analysis.

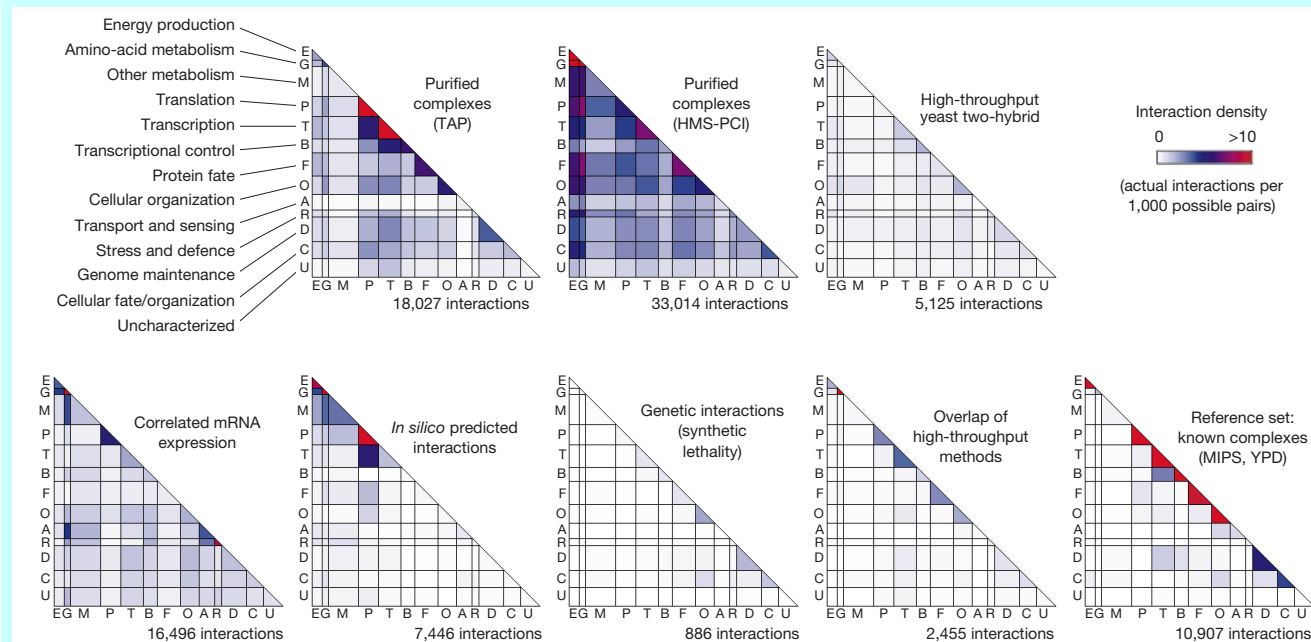


Figure 1 Large-scale interaction data and the distribution of interactions according to functional categories. Each data set is represented by a matrix showing the distribution of interactions (interaction density²³) by colour. Each axis on a matrix represents the entire yeast genome, which has been subdivided into functional categories using a catalogue of known and predicted protein functions at MIPS¹⁷. The 'uncharacterized' category is not drawn to scale, because it would encompass more than a third of each axis. Some categories were fused for conciseness, and genes annotated in multiple categories were manually assigned to one. For the large-scale purification of protein

complexes, two data sets are shown separately (one based on the TAP system¹⁰, the other based on HMS-PCI¹¹), because these are the largest to date and technical details vary considerably between them. The synthetic lethal interactions come from one initial high-throughput screen⁹ (295 interactions), but also from individual screens and experiments compiled at the MIPS database¹⁷ (note that these are derived from the literature and might thus not be entirely independent from the reference set). For this and subsequent figures, details on data sets, parameters and more examples are available in the Supplementary Information.

counting and defining interactions, and because the reference set is necessarily incomplete and may well have unknown biases itself. Nevertheless, it is evident that there are large differences between the methods and even within a method when parameters are changed. As noted previously^{12,30}, the highest accuracy is achieved for interactions supported by more than one method (Fig. 2).

There are of course many different and valid ways to count and compare interactions. In the HMS-PCI study¹¹, for example, only the interactions between the bait and the co-purified proteins were counted, not the interactions among all the proteins in a purification. We can confirm that this increases the accuracy (from 2 to 6.8% for HMS-PCI and from 12.5 to 27.8% for TAP), but it is concomitant with a strong decrease in coverage (see Supplementary Information).

An independent measure of quality is the degree to which interacting proteins are annotated with the same functional category (Fig. 1): for highly accurate data sets, a intercatiNs e ndeto gusoe on the diagonal, which shows that proteins of broadly related functions preferentially interact with each other. This correlation suggests that the interactions outside of the diagonal consist largely of false positives. We note that the reference set is particularly well clustered on the diagonal, as is the overlap of high-throughput data (Fig. 1).

Biases in interaction coverage

None of the methods covers more than 60% of the proteins in the yeast genome. Are there common biases as to which proteins are covered? We identify three areas where the high-throughput interaction data are indeed biased. First, there is a bias towards proteins of high abundance. There are no genome-wide measurements of protein abundance in yeast, but mRNA levels can be used as a crude substitute^{31,32}. A plot of interaction coverage versus mRNA abun-

dance (Fig. 3) shows that most protein interaction data sets (including the curated complexes) are heavily biased towards proteins of high abundance. However, the two genetic approaches

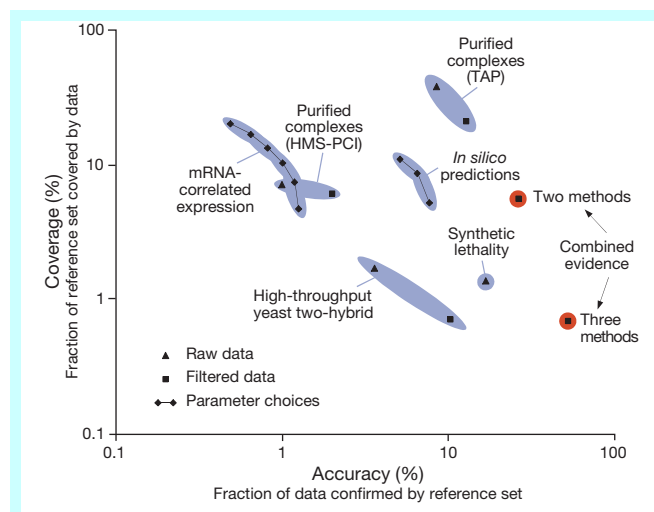


Figure 2 Quantitative comparison of interaction data sets. The various data sets are benchmarked against a reference set of 10,907 trusted interactions, which are derived from protein complexes annotated manually at MIPS¹⁷ and YPD²⁴. Coverage and accuracy are lower limits owing to incompleteness of the reference set. Each dot in the graph represents an entire interaction data set, and its position specifies coverage and accuracy (on a log-log scale). For the combined evidence, we considered only interactions supported by an agreement of two (or three) of any of the methods shown. For most data sets, raw and filtered data are shown, demonstrating the trade-off between coverage and accuracy achieved by filtering (see Supplementary Information for details on the filtering).

(two-hybrid and synthetic lethality) appear relatively unbiased. This is intriguing because these are the two methods that are, by design, largely independent of endogenous protein levels. Moreover, these two methods are especially capable of detecting transient or indirect interactions. The observed bias in the other data sets could thus reflect mainly experimental limitations, which would indicate that a large body of interactions remains undiscovered for proteins of low abundance.

Second, the data sets are biased towards particular cellular localizations of interacting proteins (Fig. 4a), for example, towards mitochondrial proteins in the case of the *in silico* predictions. As well as identifying biases, protein localization data provide an independent measure of quality for the different data sets, because proteins known to interact are usually localized similarly (Fig. 4b). Third, there is a bias in interaction coverage that relates to the degree of evolutionary novelty of proteins. We note that proteins restricted to yeast are less well covered than ancient, evolutionarily conserved proteins (see Supplementary Information).

Outlook

How many protein–protein interactions can be expected in yeast? A minimum estimate can be made by comparing the currently annotated interactions to those high-throughput interactions that are supported by more than one method. This overlap of high-throughput data is around 20 times larger than would be expected by chance (compared with randomized data sets; see Supplementary Information), which indicates a good signal-to-noise ratio. Furthermore, it consists mainly of interactions in which both partners have the same functional category and cellular localization (Figs 1 and 4). Both observations suggest that the overlap consists largely of true positives. We note, however, that less than a third of these are annotated as previously known, indicating that there should be at least three times as many interactions in yeast as there are described today. Roughly 10,000 interactions are currently known^{17,24}, which leads to a lower estimate of 30,000 interactions in yeast. The actual number, however, will probably be much higher because protein expression and interaction patterns will change during development or morphogenesis, or in response to the many

different external conditions to which yeast may be exposed in its natural environment.

Among the interactions proposed by high-throughput methods will be many false positives. In fact, we estimate that more than half of all current high-throughput data are spurious. This estimate is based on the frequent linkage of functionally unrelated proteins, often from distinct cellular compartments, which is in contrast to the reference set and the overlap data. (On average, only 21% of high-throughput interactions link proteins of the same functional category. In reality, this fraction should be higher: the overlap data has 48% and the reference set more than 80%. A similar argument can be made for the data concerning proteins from distinct cellular compartments (see Fig. 4).) For a filtered yeast two-hybrid data set, which shows medium accuracy in our benchmark (Fig. 2), the fraction of false positives has also been predicted^{8,33} to be of the order of 50%.

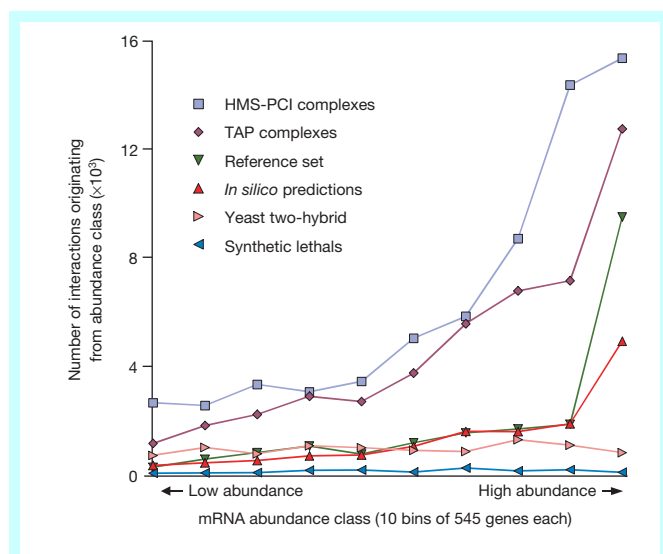


Figure 3 A bias in interaction coverage from mRNA abundance data. Using data from a survey of mRNA abundances in yeast³², we divided the yeast genome into ten mRNA abundance classes (bins) of equal size. For each data set and abundance class, we recorded the number of interactions having at least one protein in that class. Each interaction (A–B) is counted twice: once under the abundance class of partner A, and once under the abundance class of partner B.

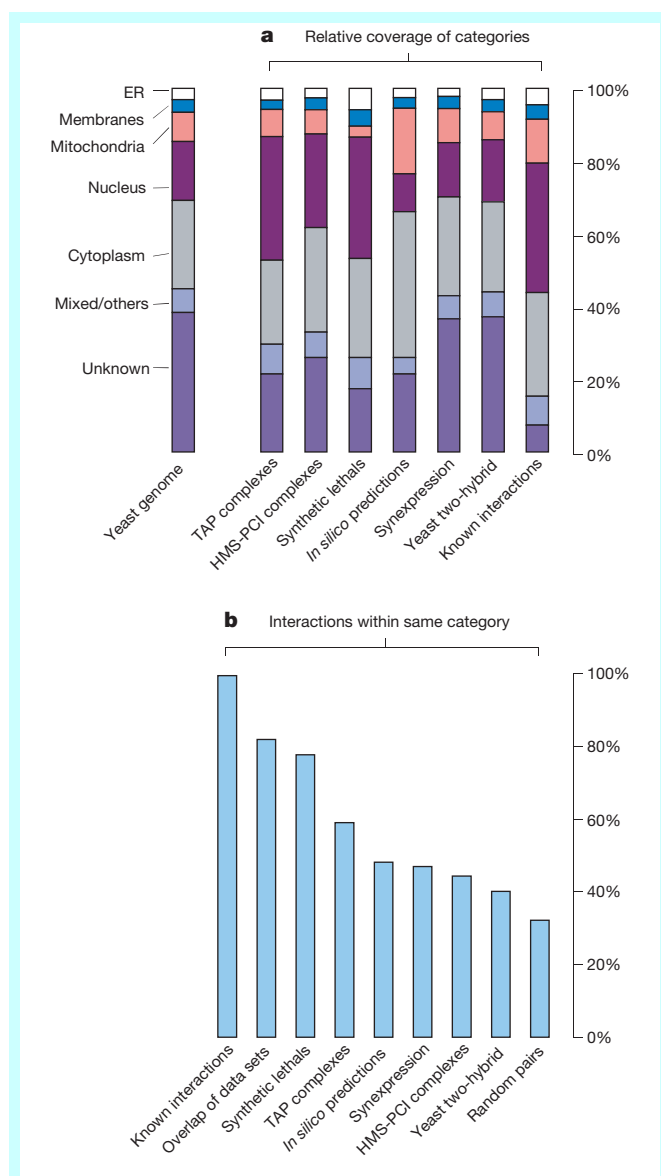


Figure 4 Protein localization and interaction coverage. Protein localizations are derived from the MIPS¹⁷ and TRIPLES (Transposon-insertion Phenotypes, Localization, and Expression in *Saccharomyces*)⁴⁰ databases. **a**, The distribution of protein localization among the proteins covered by a data set (relative coverage). ER, endoplasmic reticulum. **b**, The fraction of interactions in which both partners have the same protein localization. Here, only proteins clearly assigned to a single category are considered.

To increase coverage and to improve confidence in detected or predicted protein interactions, as many complementary methods as possible should be used, including those studied here as well as new approaches such as protein microarrays^{34,35}. Together with improvements in current technologies, transparent quality control and benchmarking, this will lead to a vast and expanding body of reliable high-throughput interaction data. Exploitation of the non-overlapping interactions will remain a challenge and will require integrated database approaches^{17,36–39} as well as careful curation and validation, as has long been good practice for individual experiments in the laboratory. □

Received 27 February; accepted 17 April 2002.

Published online 8 May 2002, DOI 10.1038/nature750.

1. Hubbard, T. *et al.* The Ensembl genome database project. *Nucleic Acids Res.* **30**, 38–41 (2002).
2. Benson, D. A. *et al.* GenBank. *Nucleic Acids Res.* **30**, 17–20 (2002).
3. Kahn, P. From genome to proteome: looking at a cell's proteins. *Science* **270**, 369–370 (1995).
4. Tucker, C. L., Gera, J. F. & Uetz, P. Towards an understanding of complex protein networks. *Trends Cell Biol.* **11**, 102–106 (2001).
5. Zamir, E. & Geiger, B. Molecular complexity and dynamics of cell-matrix adhesions. *J. Cell Sci.* **114**, 3583–3590 (2001).
6. Vidal, M. A biological atlas of functional maps. *Cell* **104**, 333–339 (2001).
7. Uetz, P. *et al.* A comprehensive analysis of protein–protein interactions in *Saccharomyces cerevisiae*. *Nature* **403**, 623–627 (2000).
8. Ito, T. *et al.* A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc. Natl Acad. Sci. USA* **98**, 4569–4574 (2001).
9. Tong, A. H. *et al.* Systematic genetic analysis with ordered arrays of yeast deletion mutants. *Science* **294**, 2364–2368 (2001).
10. Gavin, A. C. *et al.* Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415**, 141–147 (2002).
11. Ho, Y. *et al.* Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature* **415**, 180–183 (2002).
12. Marcotte, E. M., Pellegrini, M., Thompson, M. J., Yeates, T. O. & Eisenberg, D. A combined algorithm for genome-wide prediction of protein function. *Nature* **402**, 83–86 (1999).
13. Huynen, M., Snel, B., Lathe, W. 3rd & Bork, P. Predicting protein function by genomic context: quantitative evaluation and qualitative inferences. *Genome Res.* **10**, 1204–1210 (2000).
14. Overbeek, R., Fonstein, M., D'Souza, M., Pusch, G. D. & Maltsev, N. The use of gene clusters to infer functional coupling. *Proc. Natl Acad. Sci. USA* **96**, 2896–2901 (1999).
15. Cho, R. J. *et al.* A genome-wide transcriptional analysis of the mitotic cell cycle. *Mol. Cell* **2**, 65–73 (1998).
16. Hughes, T. R. *et al.* Functional discovery via a compendium of expression profiles. *Cell* **102**, 109–126 (2000).
17. Mewes, H. W. *et al.* MIPS: a database for genomes and protein sequences. *Nucleic Acids Res.* **30**, 31–34 (2002).
18. Enright, A. J., Iliopoulos, I., Kyripides, N. C. & Ouzounis, C. A. Protein interaction maps for complete genomes based on gene fusion events. *Nature* **402**, 86–90 (1999).
19. Marcotte, E. M. *et al.* Detecting protein function and protein–protein interactions from genome sequences. *Science* **285**, 751–753 (1999).
20. Dandekar, T., Snel, B., Huynen, M. & Bork, P. Conservation of gene order: a fingerprint of proteins that physically interact. *Trends Biochem. Sci.* **23**, 324–328 (1998).
21. Pellegrini, M., Marcotte, E. M., Thompson, M. J., Eisenberg, D. & Yeates, T. O. Assigning protein functions by comparative genome analysis: protein phylogenetic profiles. *Proc. Natl Acad. Sci. USA* **96**, 4285–4288 (1999).
22. Huynen, M. A. & Bork, P. Measuring genome evolution. *Proc. Natl Acad. Sci. USA* **95**, 5849–5856 (1998).
23. Ge, H., Liu, Z., Church, G. M. & Vidal, M. Correlation between transcriptome and interactome mapping data from *Saccharomyces cerevisiae*. *Nature Genet.* **29**, 482–486 (2001).
24. Costanzo, M. C. *et al.* YPD, PombePD and WormPD: model organism volumes of the BioKnowledge library, an integrated resource for protein information. *Nucleic Acids Res.* **29**, 75–79 (2001).
25. Piper, M. D., Hong, S. P., Ball, G. E. & Dawes, I. W. Regulation of the balance of one-carbon metabolism in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **275**, 30987–30995 (2000).
26. Sinclair, D. A. & Dawes, I. W. Genetics of the synthesis of serine from glycine and the utilization of glycine as sole nitrogen source by *Saccharomyces cerevisiae*. *Genetics* **140**, 1213–1222 (1995).
27. Douce, R., Bourguignon, J., Neuburger, M. & Rebeille, F. The glycine decarboxylase system: a fascinating complex. *Trends Plant Sci.* **6**, 167–176 (2001).
28. Hoffmann, R., Jung, S., Ehrmann, M. & Hofer, H. W. The *Saccharomyces cerevisiae* gene *PPH3* encodes a protein phosphatase with properties different from PPX, PP1 and PP2A. *Yeast* **10**, 567–578 (1994).
29. Kalhor, H. R., Luk, K., Ramos, A., Zobel-Thropp, P. & Clarke, S. Protein phosphatase methyltransferase 1 (Ppm1p) is the sole activity responsible for modification of the major forms of protein phosphatase 2A in yeast. *Arch. Biochem. Biophys.* **395**, 239–245 (2001).
30. Gerstein, M., Lan, N. & Jansen, R. Proteomics. Integrating interactomes. *Science* **295**, 284–287 (2002).
31. Gygi, S. P., Rochon, Y., Franza, B. R. & Aebersold, R. Correlation between protein and mRNA abundance in yeast. *Mol. Cell. Biol.* **19**, 1720–1730 (1999).
32. Holstege, F. C. *et al.* Dissecting the regulatory circuitry of a eukaryotic genome. *Cell* **95**, 717–728 (1998).
33. Mrowka, R., Patzak, A. & Herzel, H. Is there a bias in proteome research? *Genome Res.* **11**, 1971–1973 (2001).
34. Zhu, H. *et al.* Global analysis of protein activities using proteome chips. *Science* **293**, 2101–2105 (2001).
35. Zhou, H., Roy, S., Schulman, H. & Natan, M. J. Solution and chip arrays in protein profiling. *Trends Biotechnol.* **19**, S34–S39 (2001).
36. Xenarios, I. *et al.* DIP, the Database of Interacting Proteins: a research tool for studying cellular networks of protein interactions. *Nucleic Acids Res.* **30**, 303–305 (2002).
37. Mellor, J. C., Yanai, I., Clodfelter, K. H., Mintseris, J. & DeLisi, C. Predictome: a database of putative functional links between proteins. *Nucleic Acids Res.* **30**, 306–309 (2002).
38. Bader, G. D. *et al.* BIND—The Biomolecular Interaction Network Database. *Nucleic Acids Res.* **29**, 242–245 (2001).
39. Paton, N. W. *et al.* Conceptual modelling of genomic information. *Bioinformatics* **16**, 548–557 (2000).
40. Kumar, A. *et al.* Subcellular localization of the yeast proteome. *Genes Dev.* **16**, 707–719 (2002).
41. Fields, S. & Song, O. A novel genetic system to detect protein–protein interactions. *Nature* **340**, 245–246 (1989).
42. Oliver, S. Guilt-by-association goes global. *Nature* **403**, 601–603 (2000).
43. Rigaut, G. *et al.* A generic protein purification method for protein complex characterization and proteome exploration. *Nature Biotechnol.* **17**, 1030–1032 (1999).
44. Mann, M., Hendrickson, R. C. & Pandey, A. Analysis of proteins and proteomes by mass spectrometry. *Annu. Rev. Biochem.* **70**, 437–473 (2001).

Supplementary Information accompanies the paper on *Nature's* website (<http://www.nature.com>).

Acknowledgements

We thank members of the Bork group, N. Paton, A. Brass, S. Cohen, A.-C. Gavin and B. Kuster for critical discussions, and Cellzone AG for access to their interaction data before publication. C.v.M. and P.B. are supported by Bundesministerium für Bildung und Forschung, Germany, and work in Manchester is supported by the Biotechnology and Biological Sciences Research Council through its IGF (Investigating Gene Function) initiative. S.F. is supported by the National Center for Research Resources of the National Institutes of Health, and is an investigator of the Howard Hughes Medical Institute. S.F. and P.B. are supported by the Human Frontier Science Program.

Competing interests statement

The authors declare competing financial interests: details accompany the paper on *Nature's* website (<http://www.nature.com>).

Correspondence and requests for materials should be addressed to P.B. (e-mail: bork@embl-heidelberg.de).

Prokineticin 2 transmits the behavioural circadian rhythm of the suprachiasmatic nucleus

Michelle Y. Cheng^{*†}, Clayton M. Bullock^{*†}, Chuanyu Li^{*}, Alex G. Lee^{*}, Jason C. Bermak^{*}, James Belluzzi^{*}, David R. Weaver[‡], Frances M. Leslie^{*} & Qun-Yong Zhou^{*}

^{*} Department of Pharmacology, University of California, Irvine, California 92697, USA

[‡] Department of Neurobiology, University of Massachusetts Medical School, Worcester, Massachusetts 01655, USA

[†] These authors contributed equally to this work

The suprachiasmatic nucleus (SCN) controls the circadian rhythm of physiological and behavioural processes in mammals. Here we show that prokineticin 2 (PK2), a cysteine-rich secreted protein, functions as an output molecule from the SCN circadian clock. PK2 messenger RNA is rhythmically expressed in the SCN, and the phase of PK2 rhythm is responsive to light entrainment. Molecular and genetic studies have revealed that *PK2* is a gene that is controlled by a circadian clock (clock-controlled). Receptor for PK2 (PKR2) is abundantly expressed in major target nuclei of the SCN output pathway. Inhibition of nocturnal locomotor activity in rats by intracerebroventricular delivery of recombinant PK2 during subjective night, when the endogenous PK2 mRNA level is low, further supports the hypothesis that PK2 is an output molecule that transmits behavioural circadian rhythm. The high expression of PKR2 mRNA within the SCN and the positive feedback of PK2 on its own transcription through activation of PKR2 suggest that PK2 may also function locally within the SCN to synchronize output.

The organization of physiology and behaviour with recurring daily environmental conditions is an adaptation that occurs in essentially all living organisms. Circadian rhythms are regulated by three components: the circadian pacemaker, an input mechanism and an output mechanism. In mammals, the pacemaker that drives circadian rhythms resides in the suprachiasmatic nuclei (SCN) of the anterior hypothalamus¹. Environmental light/dark cycles entrain the SCN clock to the 24-h day by means of retinal-hypothalamic projections². Synchronization of SCN neurons leads to coordinated circadian outputs that regulate expressed rhythms.

A clear understanding of the molecular clock mechanisms within the SCN has emerged^{3–7}. The molecular clockwork consists of autoregulatory transcriptional and translational feedback loops that have both positive and negative elements. Positive transcriptional elements include two basic helix–loop–helix, PAS-domain-containing transcription factors, Clock and Bmal1, which form heterodimers and drive the transcription of three mammalian *period* genes (*Per1*, *Per2* and *Per3*) and two *cryptochrome* genes (*Cry1* and *Cry2*) by binding to their respective CACGTG E-box enhancers^{8,9}. The *Per* and *Cry* proteins act as negative components of the feedback loop, with *Cry* proteins having dominant inhibitory roles¹⁰. Evidence also supports the existence of an interacting molecular loop such that *Per2* positively regulates *Bmal1* transcription¹¹.

Much of our understanding of the molecular clock mechanisms results from genetic studies with mutant animals. Mice deficient in *Cry1*, *Cry2* or *Per3* show subtle changes in circadian cycle length, but retain rhythmicity^{12–15}. Mutations in *clock*, *Per1* or *Per2* genes result in a more severe circadian phenotype that includes arrhythmicity after long-term housing in constant darkness^{16–20}. Mice with disrupted *Bmal1*, or that are deficient in both *Per1* and *Per2*, exhibit the most severe phenotype: they are arrhythmic immediately after placement in constant darkness^{19–21}. Similar arrhythmicity also occurs in mice deficient in both *Cry1* and *Cry2* genes^{12,15}. The cloning of the *Tau* gene, which codes for a casein kinase ϵ and whose mutation causes a shortened circadian cycle in a strain of hamster, has revealed the importance of post-transcriptional mechanisms for

SCN clockwork²².

In contrast to SCN molecular clock mechanisms, relatively little is known about the mechanism by which circadian pacemaker systems transmit timing information to control physiology and behaviour^{6,7,23}. In *Drosophila*, pigment-dispersing factor, a neuropeptide, and *takeout*, a clock-controlled gene that seems to code for a secreted protein, are mediators that transmit circadian activity rhythms^{24,25}. In mammals, transplant studies have indicated that signals that mediate the rhythmic output of the SCN also appear to be secreted molecules^{26–28}, but their biochemical identities are unknown. Vasopressin, a neuropeptide rhythmically expressed in the SCN²⁹, may regulate rhythmic endocrine activity in the hypothalamus, pituitary and adrenal gland³⁰. However, the absence of vasopressin does not affect other SCN-regulated circadian behaviours, including locomotor activity³¹. Recently, transforming growth factor (TGF)- α has been shown to be a signalling molecule that mediates light-induced suppression of locomotion³², but with no obvious role in the locomotor rhythm in constant darkness. Albumin D-element binding protein (DBP), a transcription factor that itself is a clock-controlled gene, has also been implicated in SCN output³³. Precisely how DBP, a nuclear protein, regulates circadian output of the SCN is unclear.

We propose that an output molecule that transmits the SCN-controlled behavioural circadian rhythm should fulfil the following criteria: (1) it is a secreted molecule; (2) its production and/or release is regulated by core clock genes; (3) The production or release of this output molecule should respond to light entrainment; (4) receptor for this molecule should be expressed in primary SCN output target areas; and (5) administration of the output molecule should result in changes in circadian behaviour. We report that prokineticin 2, a recently identified secreted protein, fulfils these criteria and is shown as a SCN output molecule that transmits the circadian locomotor rhythm.

Rhythmic expression of PK2 mRNA in the SCN

We have identified two mammalian secreted proteins, prokineticin 1 (PK1) and prokineticin 2 (PK2), that potently contract prep-

arations of gastrointestinal smooth muscle strips³⁴. PK2 mRNA transcript, of approximately 1.6 kilobases (kb), is expressed in the central nervous system, as seen by northern hybridization analysis (data not shown). Further examination by *in situ* hybridization has revealed that PK2 mRNA is expressed in the SCN, among a few other discrete brain areas including the islands of Calleja, medial preoptic area of the hypothalamus, and the shell of the nucleus accumbens (Fig. 1a). To explore the time dependence of PK2 mRNA in the SCN, we entrained mice under a 12-h light/dark cycle. Quantification of PK2 mRNA in the SCN at various time points reveals a circadian oscillation profile during a light/dark cycle (Fig. 1b). PK2 mRNA is highest during the light phase (Zeitgeber time, ZT, ZT1–ZT7)—where ZT0 is light on and ZT12 is light off—and lowest during the dark phase (ZT13–ZT22). The oscillation magnitude of PK2 mRNA is very high, with the peak level at least 50-fold higher than the lowest level. PK2 mRNA is essentially undetectable in SCN during the dark phase. No oscillation of PK2 expression was observed in other positive areas such as the medial preoptic area or the islands of Calleja (Fig. 1c). It should be noted that the closely related PK1 mRNA is undetectable in the brain areas we examined, including the SCN (data not shown).

The oscillation of PK2 mRNA is maintained under constant darkness (Fig. 1b, c). After 2 days in constant darkness, PK2 mRNA levels are high from CT1–CT7 (circadian time, CT, where CT0 is subjective light on and CT12 is subjective light off), and remains low from CT13–CT22. However, after 8 days in constant darkness the period of low PK2 expression expands from about 9 h in light/dark conditions to 12 h, with a slight reduction in peak level. The robust circadian profile of PK2 mRNA in the SCN suggests that PK2 regulates SCN circadian pacemakers and/or their output.

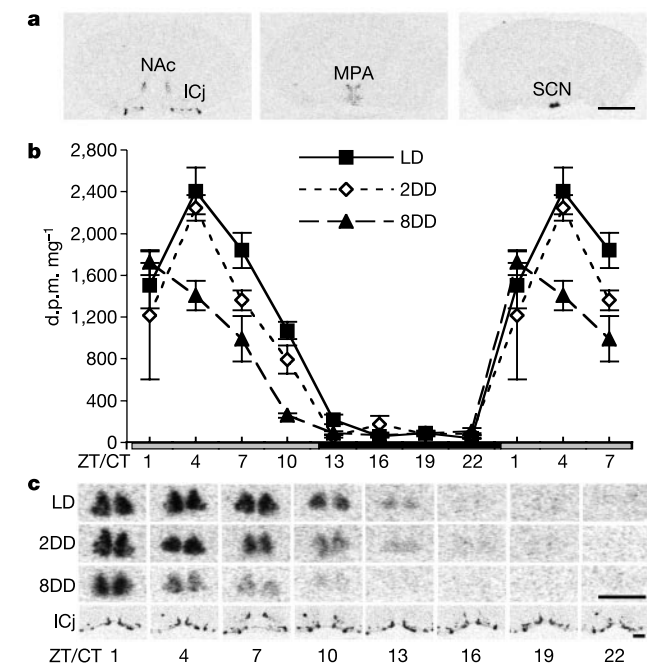


Figure 1 Rhythmic expression of PK2 mRNA in the SCN. **a**, Autoradiograms of PK2 mRNA expression in the mouse brain. NAc, nucleus accumbens; ICj, islands of Calleja; MPA, medial preoptic area; SCN, suprachiasmatic nucleus. Scale bar, 2 mm. **b**, Temporal profiles of PK2 mRNA in the SCN. Animals were studied under light/dark (LD), or constant darkness for 2 days (2DD) or 8 days (8DD). Each value is the mean $\pm$ s.e.m. of three–six animals. Data at ZT/CT1–7 are plotted twice. Shaded and filled horizontal bars indicate light and dark periods, respectively. **c**, Representative images of PK2 mRNA expression shown in **b**. Scale bar, 1 mm.

Regulation of PK2 transcription by clock genes

As E-box enhancers residing in the promoter regions are necessary for Clock–Bmal1-mediated transcriptional activation^{8,9}, we examined the putative promoter sequence of the mouse PK2 gene³⁵. Four E-box elements are found within 2.4 kb of the mouse PK2 5'-flanking region (Fig. 2a). A GenBank search reveals that these four E-box elements are conserved in the 5'-flanking sequence of the human PK2 gene, including the approximate location of each of the four E-boxes (data not shown). No E-box sequence was identified in the 5-kb 5'-flanking region of the closely related PK1 gene (data not shown).

To examine the ability of Clock–Bmal1 heterodimers to drive PK2 transcription through E-box enhancers, we cloned a 2.8-kb fragment of the mouse PK2 5'-flanking region into a promoterless

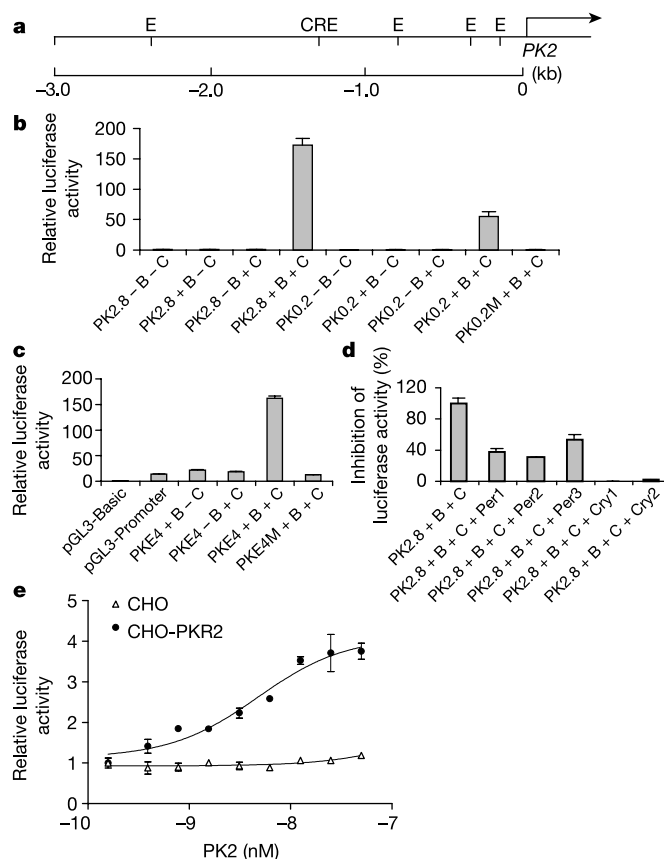


Figure 2 *In vitro* transcription analyses of the mouse PK2 gene. **a**, Location of E-boxes (E) and cAMP-responsive element (CRE) within the 5'-flanking region of the mouse PK2 gene. Numbers represent distance in kilobases from the putative transcription start site, marked as 0. **b**, Transcriptional activation of the luciferase reporter containing 2.8 kb (PK2.8-Luc) and 200 bp (PK0.2-Luc) of the 5'-flanking region of the mouse PK2 gene. PK0.2M-Luc represents a mutated E-box (GGATCT). C, B, indicate Clock and Bmal1, respectively. Luciferase activity of PK2.8-Luc in the absence of Clock and Bmal1 was designated as 1. **c**, Transcriptional activation of SV40-driven luciferase reporter (pGL3-Promoter) containing 72 bp of all four E-boxes and their immediate flanking sequences (PK4E-Luc). All four E-boxes were mutated in PK4EM-Luc. Luciferase activity of pGL3-Basic was designated as 1. **d**, Inhibition of Clock–Bmal1-mediated transcription from PK2.8-Luc by Per proteins and Cry proteins. Luciferase activity of PK2.8-Luc in the presence of Clock (C) and Bmal1 (B) was designated as 100%. **e**, Positive effect of PKR2 activation on Clock–Bmal1-mediated transcription from PK2.8-Luc. Chinese hamster ovary (CHO) cells and CHO cells that stably express PKR2 (CHO-PKR2) were transfected with PK2.8-Luc, along with Clock and Bmal1. Luciferase activity of PK2.8-Luc in CHO cells in the absence of PK2 treatment was designated as 1. Each value shown (**b–e**) is the mean $\pm$ s.e.m. of three replicates from a single assay. The results shown were representative of at least three independent experiments.

luciferase reporter vector (pGL3-Basic). Clock and Bmal1 together, but neither alone, produced a strong increase in transcriptional activity (172-fold, Fig. 2b). One proximal E-box, residing within 130 bp upstream of the putative PK2 transcriptional start site, was also transactivated by Clock–Bmal1 heterodimers (Fig. 2b). This Clock–Bmal1-dependent activation was abolished when the E-box CACGTG was mutated. We also examined the transcriptional activity of a 72-bp construct in which the four E-boxes and their immediate flanking sequence were linked together and ligated into a luciferase reporter vector containing the SV40 promoter (pGL3-Promoter). Clock and Bmal1 together (but neither alone) caused a substantial increase in transcription through these four E-boxes (162-fold, Fig. 2c). This Clock–Bmal1-dependent activation was reduced to background levels when all four E-boxes were mutated. These results indicate that transcription by Clock–Bmal1 heterodimers requires at least one E-box. Using the 2.8-kb PK2 promoter, we next examined the effects of negative elements of the SCN clockwork. Mouse period proteins caused 40–60% inhibition of Clock–Bmal1-induced transcription, whereas there was almost complete inhibition mediated by Cry1 and Cry2 (Fig. 2d). Taken together, these *in vitro* studies indicate that clock gene products can regulate PK2 transcription.

Altered PK2 rhythm in molecular clockwork mutant mice

We next examined PK2 mRNA levels in the SCN of mutant mice deficient in Clock (*Clk*^{−/−})¹⁷ or cryptochromes (*Cry1*^{−/−}, *Cry2*^{−/−}; Cry-deficient)¹⁵. The circadian rhythm of PK2 mRNA expression in the SCN is severely blunted in *Clk*^{−/−} mice under light/dark conditions (Fig. 3a). PK2 mRNA is detectable only at ZT5–ZT8. The reduced peak level of PK2 mRNA in *Clk*^{−/−} mice is also shifted (from about ZT4 to ZT8). It should also be noted that PK2 mRNA levels in the medial preoptic area did not change in *Clk*^{−/−} mice. These results show that the decrease in PK2 gene expression in *Clk*^{−/−} mice is specific for the SCN circadian clock. The severe blunting of PK2 circadian rhythm in *Clk*^{−/−} mice indicates that clock has a positive regulatory effect on the expression of PK2.

We also examined PK2 mRNA levels in the SCN of Cry-deficient mice. In contrast to the oscillation pattern seen in the SCN of wild-type mice, PK2 mRNA levels were very low at both CT6 and CT18 in

Cry-deficient mice (Fig. 3b). PK2 mRNA levels in the medial preoptic area were not observably different compared to wild-type mice at either circadian time. The apparent low levels of PK2 mRNA at both circadian times in the SCN of Cry-deficient mice seems unexpected as Cry proteins are major inhibitors of Clock–Bmal1-mediated transcription¹⁰, and tonic mid-high Per1 and Per2 mRNA levels have been observed in Cry-deficient mice^{12,36}. The immediate arrhythmicity of Cry-deficient mice in constant darkness indicates the disruption of the functional circadian clock¹⁵. In the absence of functional positive and negative limbs of the feedback loops, the absolute levels of PK2 mRNA probably depend on basal promoter activity and RNA stability. The essentially undetectable levels of PK2 mRNA in the SCN during dark phase (Fig. 1b, c) suggest low basal PK2 promoter activity. It should also be noted that there are multiple copies of mRNA destability signal (AUUUA) in the 3′ untranslated regions of both the human and mouse PK2 mRNAs. These combined effects may contribute to the low levels of PK2 mRNA in the SCN of Cry-deficient mice. Indeed, simulating the data from the light-induced PK2 mRNA expression (Fig. 4a) generated a half-life of PK2 mRNA of a mere 36 min³⁷. Taken together, these genetic studies support the hypothesis that PK2 is a clock-controlled gene.

Response of PK2 rhythm to light entrainment

We next examined PK2 mRNA levels in the SCN in response to light pulses under conditions of constant darkness. PK2 mRNA increased rapidly and transiently 30 min after light exposure at both CT14 and CT23 (Fig. 4a, b). At CT14, PK2 mRNA remained undetectable without light pulse. After exposure to light, PK2 mRNA increased

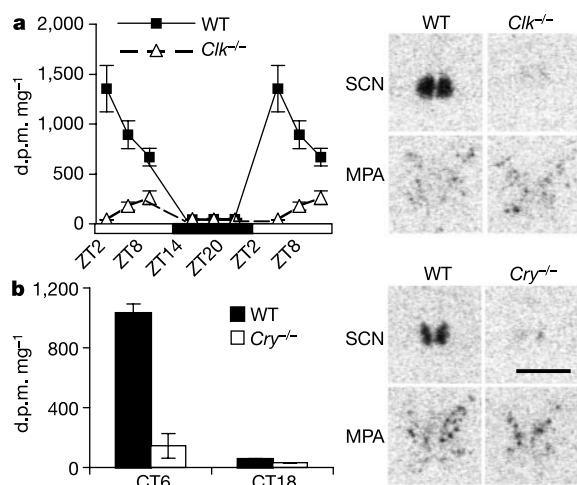


Figure 3 Rhythmic expression of PK2 mRNA in Clock-deficient (*Clk*^{−/−}) or cryptochrome-deficient (*Cry*^{−/−}) mice. **a**, Temporal profiles of PK2 mRNA in the SCN of *Clk*^{−/−} mice. Representative autoradiograms of coronal sections at ZT2 of wild-type (WT) and *Clk*^{−/−} mice in the SCN and medial preoptic area (MPA) are shown on the right. **b**, Expression of PK2 mRNA in the SCN of *Cry*^{−/−} mice at CT6 and CT18. Representative autoradiograms at CT6 of wild-type and *Cry*^{−/−} mice in the SCN and MPA are shown on the right. Scale bar, 1 mm. Each value is the mean ± s.e.m. of five–six animals.

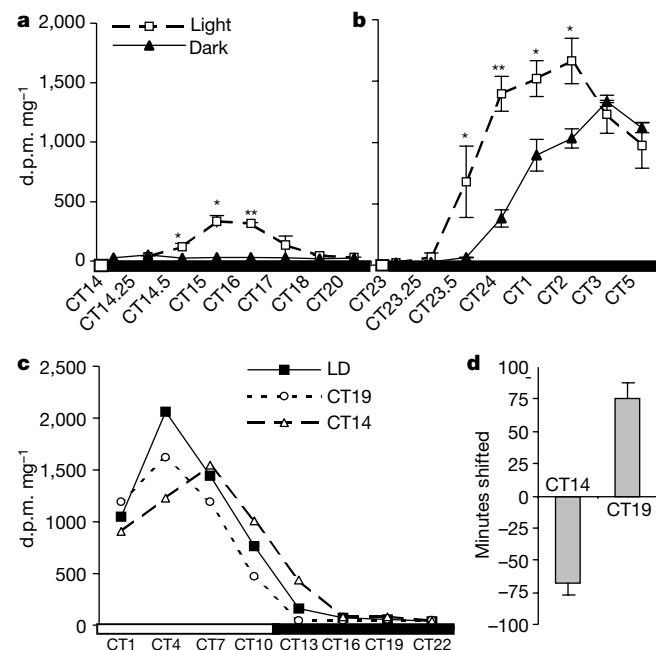


Figure 4 PK2 rhythm in the SCN responds to light entrainment. Animals were exposed to a 15-min light pulse at CT14 (**a**) or CT23 (**b**) on the first day in constant darkness. Open symbols indicate light-exposed animals; filled symbols indicate time-matched controls maintained in constant darkness. Each value is the mean ± s.e.m. of four–six animals. Asterisk, *P* < 0.05; double asterisk, *P* < 0.01 (Student's *t*-tests). **c**, Light pulses shift PK2 rhythm in the SCN. Animals were exposed to a 1-h light pulse at CT14–15 or CT19–20 on the first day in constant darkness. Controls received no light pulse. Shaded and filled horizontal bars indicate light and dark periods under conditions of constant darkness, respectively. One of three experiments is shown. **d**, Quantification of phase shifts (in minutes, *n* = 3) shown in **c**. The shift was measured at the half-peak level of PK2 mRNA.

after 30 min, peaked at 1–2 h, and returned to control levels 4 h after exposure. At CT23, the level of PK2 mRNA rose with or without light pulse; however, in the presence of a light pulse, PK2 mRNA increased more rapidly and robustly, and returned to control levels 4 h after exposure. At CT3, however, light pulse had no enhancing effect on PK2 mRNA levels (data not shown). The light-inducibility of PK2 mRNA may result from activation of a putative cyclic AMP-responsive element (CRE) in the PK2 promoter (Fig. 2a).

We also examined whether light pulses that shift the phases of SCN-controlled behavioural rhythms also shift the phase of PK2 rhythm. Light pulses delivered during early subjective night cause phase delay of locomotor rhythm, whereas light pulses delivered during late subjective night cause phase advance of locomotor rhythm^{38–40}. Light pulses administered during early subjective night (CT14) cause a delay of about 70 min in the phase of PK2 mRNA expression (Fig. 4c, d). Similarly, light pulses administered during late subjective night (CT19) cause about a 75-min advance in the phase of PK2 mRNA expression. These results indicate that light pulses that shift the phase of SCN-controlled behavioural rhythm alter the phase of PK2 expression. These light-induced shifts of PK2 rhythm correlate well with the shifts of locomotor rhythm.

Expression of PK2 receptor in SCN output targets

Recently, we have cloned two closely related receptors for PK1 and PK2. These receptors, named prokineticin receptor 1 and 2 (PKR1 and PKR2), belong to the family of G-protein-coupled receptors (C.M.B., M.Y.C. and Q.-Y.Z., manuscript in preparation). *In situ* hybridization revealed that PKR2 mRNA is expressed in a few discrete areas in the mouse brain. PKR2 mRNA is abundantly or moderately expressed in several brain regions, including the paraventricular hypothalamic nucleus (PVN), dorsal medial hypothalamic nucleus (DMH), paraventricular and paratenial thalamic nuclei (PVT/PT), paracentral thalamic nucleus (PC), lateral habenular nucleus (LHb), lateral septal nucleus (LS), lateral globus pallidus (LGP) and amygdala (Fig. 5; see also Supplementary Information Fig. 1). The highest expression of PKR2 was found in PVT. Notably, many of the nuclei where PKR2 mRNA expression occurs are primary target areas of SCN efferents^{1,2,41–44}. The high expression of PKR2 mRNA in these primary target areas of SCN output suggests that PK2 may serve as a signalling molecule mediating SCN output.

PKR2 mRNA is also extensively expressed in the SCN (Fig. 5c, h,

m). In contrast to the robust rhythm of PK2 mRNA expression, PKR2 mRNA levels in the SCN and other regions are non-oscillating (data not shown). High-magnification microscopic images show that most neurons within the SCN express both PK2 and PKR2 mRNAs (Fig. 5k–n). This co-expression of PK2 and PKR2 mRNAs in the SCN suggests that the PK2/PKR2 system may have a function in synchronizing SCN output. We tested the possibility of positive feedback of PKR2 activation on PK2 gene expression by stably expressing PKR2 in Chinese hamster ovary cells (CHO-PKR2) and transfecting CHO-PKR2 cells with the 2.8-kb PK2 promoter fused to the luciferase reporter gene. Activation of PKR2 by PK2 stimulates Clock–Bmal1-mediated transcription of the 2.8-kb PK2 promoter in a dose-dependent manner (4.0 ± 0.4 fold, mean $\pm$ s.e.m., $n = 3$, Fig. 2e). Effector concentration for half-maximum response (EC_{50}) values of PK2 were 7.8 ± 2.9 nM ($n = 3$), similar to EC_{50} values (7.3 ± 1.5 nM, $n = 4$) obtained from the calcium mobilization of PKR2 activation by PK2 (C.M.B., M.Y.C. and Q.-Y.Z., manuscript in preparation). Further experiments indicate that the stimulatory effects of PKR2 activation on the activity of the 2.8-kb PK2 promoter depend on clock genes. In the absence of Clock and Bmal1, activation of PKR2 alone by PK2 treatment (50 nM for 6 h) did not activate transcription of the 2.8-kb PK2 promoter (data not shown). In the presence of Clock and Bmal1, the stimulatory effect of PKR2 activation on the transcription of the 2.8-kb PK2 promoter was abolished by Cry1 and Cry2 (data not shown). Thus, PK2 may activate its own transcription in the SCN by means of activation of PKR2 in a manner dependent on the products of clock genes.

PK2 administration alters circadian locomotor activity

To assess directly whether PK2 mediates the circadian behavioural output in rats, we studied the effects of intracerebroventricular (ICV) injection of recombinant PK2 on wheel-running behaviour. Figure 6 shows that PK2 delivery suppressed the high level of wheel-running behaviour associated with dark phase, whereas administration of saline only slightly inhibited wheel-running behaviour. In contrast to the normally quiescent daytime activity, PK2-treated rats were active during the next subjective day. PK2 infusion caused a slight delay of wheel-running rhythm in subsequent days; however, this delay was not significantly different from controls. These results indicate that PK2, which is normally at high levels during subjective day, inhibits wheel-running activity. During subjective

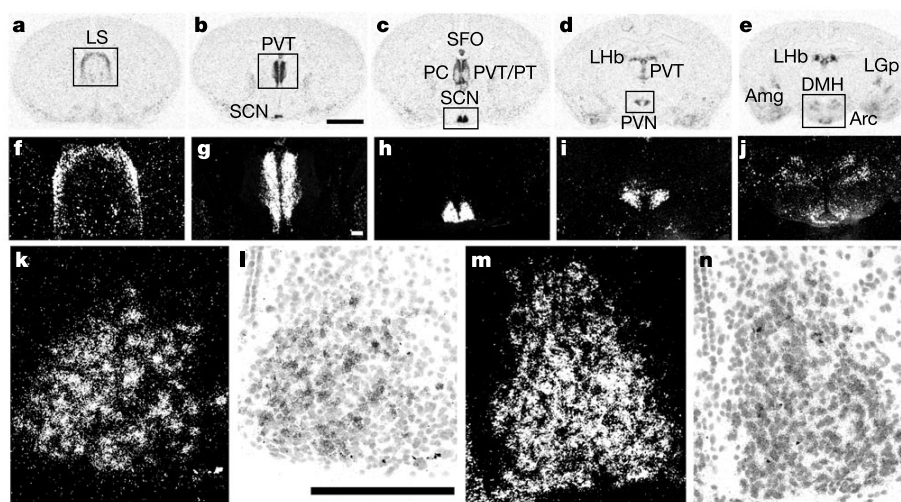


Figure 5 Expression of PK2 receptor (PKR2) mRNA in mouse brain. **a–e**, Autoradiograms of PKR2 mRNA in the lateral septum (LS), paraventricular thalamic (PVT) and hypothalamic nucleus (PVN), suprachiasmatic nucleus (SCN), paratenial nucleus (PT), subfornical organ (SFO), paracentral nucleus (PC), lateral habenula (LHb), amygdala (Amg), lateral globus pallidus (LGP), dorsal medial hypothalamic nucleus (DMH) and arcuate nucleus (Arc).

Scale bar, 2 mm. **f–j**, Dark-field images of PKR2 mRNA from boxed regions in **a–e**, respectively. **k–n**, High-magnification images of PK2 (**k, l**) and PKR2 mRNA (**m, n**) in the SCN. Both dark field images (**k, m**) and sections stained with cresyl violet (**l, n**) are shown. Scale bar, 0.5 mm (**f–n**). All images are obtained from mouse brains collected at ZT1.

night, when endogenous PK2 mRNA levels are low, the wheel-running activity is disinhibited. The increased wheel-running activity observed on the next subjective day in PK2-treated rats may result from desensitization of the PKR2 receptor system caused by PK2 administration. Thus, PK2 is an output molecule that transmits the circadian locomotor rhythm of the SCN clock.

Discussion

We show that expression of PK2 mRNA in the SCN oscillates in a circadian fashion under a normal light/dark cycle as well as in constant darkness. The oscillation amplitude of PK2 mRNA in the SCN is very high, with peak levels (day-time) and trough levels (night-time) differing by at least 50-fold. The magnitude of PK2 oscillation in the SCN is much higher than that of other known clock or clock-controlled genes, including *Per* genes, *Cry* genes and *Bmal1*^{6,45–48}. The positive feedback of PK2 on its own transcription may contribute to its high oscillation magnitude. Both *in vitro* transcription assays and analyses of mutant mice deficient in *cryptochrome* genes and *clock* support the hypothesis that PK2 is a clock-controlled gene. PK2 transcription is activated by Clock and Bmal1 heterodimers through E-box enhancers in the 5'-flanking region, and this transactivation is also inhibited by members of the negative limb of the central molecular clockwork, including *Per1*, *Per2*, *Per3*, *Cry1* and *Cry2*. The organization of E-box enhancers in the 5'-flanking region and the oscillation profile of PK2 in the SCN are most similar to those of *Per1* (refs 45, 46). Light pulse experiments show that PK2 rhythm in the SCN responds to light entrainment. Similar to *Per1* and *Per2* (refs 45, 46, 48), PK2 responds to light induction only during subjective night. Light pulses delivered during early or late night result in delay or advance

of PK2 rhythm in SCN, respectively. These light-induced changes in PK2 rhythm are consistent with the shifts of behavioural rhythm^{37–39}.

PK2 delivery experiments provide further pharmacological evidence that PK2 is an output molecule that transmits the locomotor activity rhythm of the SCN. Intracerebroventricular administration of PK2 during subjective night suppresses the high level of nocturnal wheel-running behaviour in rats. These behavioural studies are consistent with the distribution of PKR2 mRNA in major primary target areas of SCN output pathways, including PVT/PT, PVN, DMH, LHb and LS^{1,2,41–44}. Of the known SCN targets, one area that is devoid of PKR2 mRNA is the subparaventricular zone of the hypothalamus (sPVZ)^{42,43}. The SCN innervation of sPVZ probably represents the passage of SCN fibres that actually project to thalamic and/or hypothalamic targets such as PVT and DMH, which express PKR2 (ref. 43). It is also possible that other SCN-derived locomotor inhibitory or activating signalling molecules target this region⁴⁹. Recently, it has been reported that receptor for TGF- α is expressed in sPVZ and that infusion of TGF- α inhibits nocturnal wheel-running behaviour³². It is possible that PK2 and TGF- α may represent complementary SCN signalling molecules that transmit locomotor rhythm, with one responding to the circadian clock and the other to light. Another hypothesis is that TGF- α is a downstream signalling molecule of PK2 such that its release from SCN neurons targeting sPVZ is regulated by the PK2/PKR2 system, which is under the control of central molecular clockwork. The widespread expression of PK2 in SCN neurons and PKR2 in most known SCN targets suggests that, in addition to locomotor rhythm, PK2 may also regulate other SCN-controlled circadian rhythms including the sleep/wake, feeding and endocrine rhythms, that is, PK2 may signal 'circadian day'. It will be interesting to examine the role of PK2 in these SCN-controlled circadian rhythms, as well as the involvement of PK2/PKR2 signalling within SCN in synchronizing a coherent output. □

Methods

Animals

We used male adult C57Bl6 mice (Taconic Farms) for all experiments, unless otherwise indicated. Animals were entrained to standard 12 h light/dark cycles for two weeks, then kept either under light/dark or constant darkness conditions. Sections through the SCN of *Clk*^{-/-} mice (on a BALB/c genetic background) and of *Cry*-deficient mice (*Cry1*^{-/-} and *Cry2*^{-/-}; on a C57BL/6-129 hybrid background) and their respective littermate controls were obtained as previously described^{11,29}. Mice were subjected to light pulses at 400 lx for 15 min (Fig. 4a, b) and 400 lx for 1 h (Fig. 4c).

In situ hybridization

We generated antisense and sense riboprobes containing the coding region of mouse PK2 (GenBank accession number AF487280; residues 1–528) or the 3' untranslated region (UTR) of mouse PKR2 (GenBank accession number AF487279; residues 1147–2211). The 3' UTR of PKR1 (AF487278) and PKR2 were used, as these receptors are over 80% identical at the nucleic-acid sequence level in their coding regions. We used 20- μ m sections for all experiments except the studies shown in Fig. 3 and Fig. 4a, b (where sections were 15 μ m). We processed *in situ* hybridizations as described⁵⁰. We analysed the mRNA distributions in autoradiograms and emulsion-dipped sections. Specific hybridization signals were quantitatively analysed using a video-based computer image analysis system (MCID, Imaging Research). A calibration curve of optical density versus radioactivity (disintegrations per minute (d.p.m.) per mg tissue wet weight) was constructed using ¹⁴C standards. Specific hybridization signals in SCN were obtained by subtracting background values obtained from adjacent brain areas that have no hybridization signal. Data were normalized with respect to the differences between signal intensities in equal areas of SCN.

Transcriptional assay

Human embryonic kidney (HEK293) and NIH3T3 cells were grown in DMEM supplemented with 10% fetal bovine serum. Cells were plated and transfected using lipofectamine (Invitrogen) in six-well plates. Ten nanograms of firefly luciferase reporter plasmid, 250 ng of mouse Bmal1 and mouse Clock and 500 ng *Per1*, *Per2*, *Per3*, *Cry1* or *Cry2* were used^{8–11,29}. The constant amount of DNA (1 μ g per well) was adjusted with pCDNA3.1 vector as a carrier. Forty-eight hours after transfection, cells were washed and lysed with 200 μ l reporter lysis buffer. Twenty microlitres of cell extract was mixed with Luciferase Assay Reagent (Promega) and the reaction monitored for 10 s in a Monolight 2010 luminometer (Analytical Luminescence Laboratory). For CHO cells and CHO cells that stably express PKR2 (CHO-PKR2), cells were grown in α -MEM, and transfected with the same method as for HEK293 and NIH3T3 cells. Forty-two hours after transfection,

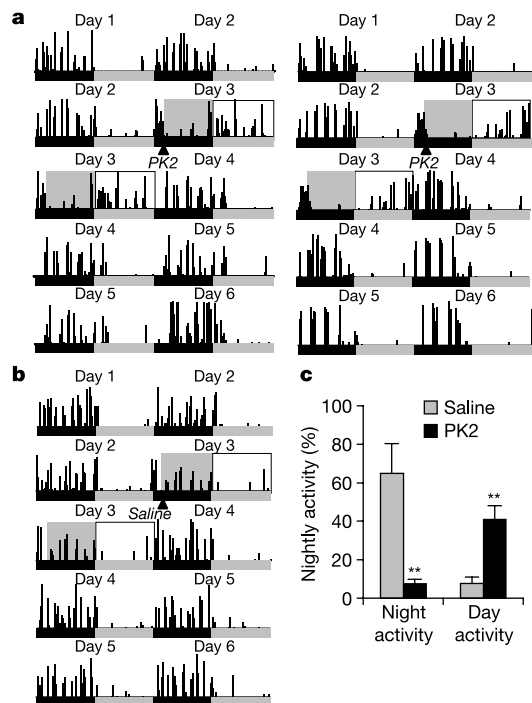


Figure 6 Effects of ICV delivery of recombinant human PK2 on wheel-running activity in rats. **a**, **b**, Representative actograms of rats injected with PK2 (**a**) or saline (**b**) at CT14. Shaded and filled horizontal bars indicate subjective light and dark periods under constant darkness, respectively. Highlighted areas indicate night locomotor activity after treatment. Boxed areas represent locomotor activity during subjective day after treatment. **c**, Quantification of night and day locomotor activity after delivery of PK2 ($n = 9$) or saline ($n = 8$). The effect of PK2 on wheel-running activity during the night phase (CT15–CT24) and day phase (CT0–CT12) was expressed as a percentage of pre-test nightly activity. Double asterisk, $P < 0.01$ (Student's *t*-test).

cells were treated with increasing concentrations of purified recombinant human PK2 (0–50 nM) for 6 h and then lysed. Luciferase activities were normalized to protein concentration.

Reporter constructs were made as follows: a 2.8-kb 5′-flanking region of the mouse PK2 gene (provided by G. Kreil and A. Jilek) was subcloned into a pGL3-Basic vector to generate PK2.8-Luc. A 170-bp 5′-flanking region of the mouse PK2 gene containing a single E-box was synthesized with four long oligonucleotides and subcloned into the pGL3-Basic vector to generate PK0.2-Luc. PK0.2M-Luc, with the E-box CACGTG mutated, was created similarly with mutant oligonucleotides. PK4E-Luc was made by subcloning annealed 72-bp oligonucleotides, which contain the four copies of E-box in the 5′-flanking region of the mouse PK2 gene linked in tandem, into the pGL3-promoter vector. A comparable construct with all four E-boxes mutated (PK4EM-Luc) was similarly constructed with mutant oligonucleotides.

Intracerebroventricular delivery of recombinant PK2

Recombinant human PK2 was produced and purified as described previously³⁴. Male Sprague–Dawley rats (300–350 g; Charles River Laboratory) were used to assess pharmacological effects of PK2 on locomotor rhythms. We implanted a 23-gauge guide cannula into the lateral ventricle (0.4 mm rostral to bregma, 3.0 mm ventral to dura, 2.4 mm lateral to midline). A 30-gauge stylet was placed in the guide cannula to maintain patency. Rats were housed in cages containing a running wheel for two weeks in light/dark conditions, placed in complete darkness for 2–5 days, and then subjected to ICV injection of recombinant PK2 at CT14 (under dim red light <1 lu). During ICV injections, the stylet was removed and a 30-gauge injector, attached to a Hamilton syringe by plastic tubing, was inserted. Each animal received either 5 µl of PK2 (1 µg µl^{−1}) or saline over a 2-min period. Animals were immediately returned to their cages and wheel-running activities were monitored in constant darkness with a PC system (Colbourn Instrument). Running-wheel data were analysed with in-house software at intervals of 5 min. Inhibition of nocturnal wheel-running behaviour was measured as activity counts during CT15–CT24, and expressed as a percentage of average counts obtained from the same time intervals of two days before treatment. Activation of wheel-running behaviour during subjective day in PK2-treated rats was measured as activity counts during CT0–CT12, and expressed as a percentage of average counts obtained from two nights (CT13–CT24) before treatment. Vasoactive intestinal peptide, an unrelated neuropeptide expressed in the SCN, behaved similarly to saline.

Received 4 January; accepted 15 April 2002.

- Klein, D. C., Moore, R. Y. & Reppert, S. M. (eds) *Suprachiasmatic Nucleus: The Mind's Clock* 467 (Oxford Univ. Press, New York, 1991).
- Moore, R. Y. Circadian rhythms: basic neurobiology and clinical applications. *Ann. Rev. Med.* **48**, 253–266 (1997).
- Dunlap, J. C. Molecular bases for circadian clocks. *Cell* **96**, 271–290 (1999).
- Gillette, M. U. & Tischkau, S. A. Suprachiasmatic nucleus: the brain's clock. *Rec. Prog. Hormone Res.* **54**, 33–59 (1999).
- Hastings, M. H. Circadian clockwork: two loops are better than one. *Nature Rev. Neurosci.* **1**, 143–146 (2000).
- Reppert, S. M. & Weaver, D. R. Molecular analysis of mammalian circadian rhythms. *Ann. Rev. Physiol.* **63**, 647–676 (2001).
- Allada, R., Emery, P., Takahashi, J. S. & Rosbash, M. Stopping time: the genetics of fly and mouse circadian clocks. *Ann. Rev. Neurosci.* **24**, 1091–1119 (2001).
- Gekakis, N. *et al.* Role of the CLOCK protein in the mammalian circadian mechanism. *Science* **280**, 1564–1569 (1998).
- Hogenesch, J. B., Gu, Y. Z., Jain, S. & Bradfield, C. A. The basic-helix-loop-helix-PAS orphan MOP3 forms transcriptionally active complexes with circadian and hypoxia factors. *Proc. Natl Acad. Sci. USA* **95**, 5474–5479 (1998).
- Kume, K. *et al.* mCRY1 and mCRY2 are essential components of the negative limb of the circadian clock feedback loop. *Cell* **98**, 193–205 (1999).
- Shearman, L. P. *et al.* Interacting molecular loops in the mammalian circadian clock. *Science* **288**, 917–924 (2000).
- Vitaterna, M. H. *et al.* Differential regulation of mammalian period genes and circadian rhythmicity by cryptochromes 1 and 2. *Proc. Natl Acad. Sci. USA* **96**, 12114–12119 (1999).
- Thresher, R. J. *et al.* Role of mouse cryptochrome blue-light photoreceptor in circadian photoreponses. *Science* **282**, 1490–1494 (1998).
- Shearman, L. P., Jin, X., Lee, C., Reppert, S. M. & Weaver, D. R. Targeted disruption of the mPer3 gene: subtle effects on circadian clock function. *Mol. Cell Biol.* **20**, 6269–6275 (2000).
- van der Horst, G. T. J. *et al.* Mammalian Cry1 and Cry2 are essential for maintenance of circadian rhythms. *Nature* **398**, 627–630 (1999).
- Cermakian, N., Monaco, L., Pando, M. P., Dierich, A. & Sassone-Corsi, P. Altered behavioural rhythms and clock gene expression in mice with a targeted mutation in the Period1 gene. *EMBO J.* **20**, 3967–3974 (2001).
- Vitaterna, M. H. *et al.* Mutagenesis and mapping of a mouse gene, *Clock*, essential for circadian behaviour. *Science* **264**, 719–725 (1994).
- King, D. P. *et al.* Positional cloning of the mouse circadian *Clock* gene. *Cell* **89**, 641–653 (1997).
- Bae, K. *et al.* Differential functions of mPer1, mPer2, and mPer3 in the SCN circadian clock. *Neuron* **30**, 525–536 (2001).
- Zheng, B. *et al.* Nonredundant roles of the mPer1 and mPer2 genes in the mammalian circadian clock. *Cell* **105**, 683–694 (2001).
- Bunger, M. K. *et al.* Mop3 is an essential component of the master circadian pacemaker in mammals. *Cell* **103**, 1009–1017 (2000).
- Lowrey, P. L. *et al.* Positional syntenic cloning and functional characterization of the mammalian circadian mutation *tau*. *Science* **288**, 483–491 (2000).
- Hardin, P. E. From biological clock to biological rhythms. *Genome Biol.* **1**, 1023.1–1023.5 (2000).

- Renn, S. C., Park, J. H., Rosbash, M., Hall, J. C. & Taghert, P. H. A pdf neuropeptide gene mutation and ablation of PDF neurons each cause severe abnormalities of behavioural circadian rhythms in *Drosophila*. *Cell* **99**, 791–802 (1999).
- Sarov-Blat, L., So, W. V., Liu, L. & Rosbash, M. The *Drosophila* takeout gene is a novel molecular link between circadian rhythms and feeding behaviour. *Cell* **101**, 647–656 (2000).
- Ralph, M. R., Foster, R. G., Davis, F. C. & Menaker, M. Transplanted suprachiasmatic nucleus determines circadian period. *Science* **247**, 975–978 (1990).
- Silver, R., LeSauter, J., Treco, P. A. & Lehman, M. N. A diffusible coupling signal from the transplanted suprachiasmatic nucleus controlling circadian locomotor rhythms. *Nature* **382**, 810–813 (1996).
- Earnest, D. J., Liang, F. Q., Ratcliff, M. & Cassone, V. M. Immortal time: circadian clock properties of rat suprachiasmatic cell lines. *Science* **283**, 693–695 (1996).
- Jin, X. *et al.* A molecular mechanism regulating rhythmic output from the suprachiasmatic circadian clock. *Cell* **96**, 57–68 (1999).
- Kalsbeek, A., van Heerikhuizen, J. J., Wortel, J. & Buijs, R. M. A diurnal rhythm of stimulatory input to the hypothalamo-pituitary-adrenal system as revealed by timed intrahypothalamic administration of the vasopressin V1 antagonist. *J. Neurosci.* **16**, 5555–5565 (1996).
- Boer, G. J., van Esseveldt, K. E., van der Geest, B. A., Duindam, H. & Reitveld, W. J. Vasopressin-deficient suprachiasmatic nucleus grafts reinstate circadian rhythmicity in suprachiasmatic nucleus-lesioned arrhythmic rats. *Neuroscience* **89**, 375–385 (1999).
- Kramer, A. *et al.* Regulation of daily locomotor activity and sleep by hypothalamic EGF receptor signalling. *Science* **294**, 2511–2515 (2001).
- Lopez-Molina, L., Conquet, F., Dubois-Dauphin, M. & Schibler, U. The DBP gene is expressed according to a circadian rhythm in the suprachiasmatic nucleus and influences circadian behaviour. *EMBO J.* **16**, 6762–6771 (1997).
- Li, M., Bullock, C. M., Knauer, D. J., Ehlert, F. J. & Zhou, Q. Y. Identification of two prokineticin cDNAs: recombinant proteins potently contract gastrointestinal smooth muscle. *Mol. Pharmacol.* **59**, 692–698 (2001).
- Jilek, A., Engel, E., Beier, D. & Lepperdinger, G. Murine Bv8 gene maps near a synteny breakpoint of mouse chromosome 6 and human 3p21. *Gene* **256**, 189–195 (2000).
- Okamura, H. *et al.* Photic induction of *mPer1* and *mPer2* in *Cry*-deficient mice lacking a biological clock. *Science* **286**, 2531–2534 (1999).
- Wilsbacher, L. D. *et al.* Photic and circadian expression of luciferase in mPeriod1-luc transgenic mice *in vivo*. *Proc. Natl Acad. Sci. USA* **99**, 489–494 (2002).
- Daan, S. & Pittendrigh, C. S. A functional analysis of circadian pacemakers in nocturnal rodents. *J. Comp. Physiol.* **106**, 253–266 (1976).
- Pittendrigh, C. S. Temporal organization: reflections of a Darwinian clock-watcher. *Ann. Rev. Physiol.* **55**, 16–54 (1993).
- Roenneberg, T. & Foster, R. G. Twilight times: light and the circadian system. *Photochem. Photobiol.* **66**, 549–561 (1997).
- Watts, A. G., Swanson, L. W. & Sanchez-Watts, G. Efferent projections of the suprachiasmatic nucleus: I. Studies using anterograde transport of *Phaseolus vulgaris* leucoagglutinin in the rat. *J. Comp. Neurol.* **258**, 204–229 (1987).
- Watts, A. G. & Swanson, L. W. Efferent projections of the suprachiasmatic nucleus: II. Studies using retrograde transport of fluorescent dyes and simultaneous peptide immunohistochemistry in the rat. *J. Comp. Neurol.* **258**, 230–252 (1987).
- Leak, R. K. & Moore, R. Y. Topographic organization of suprachiasmatic nucleus projection neurons. *J. Comp. Neurol.* **433**, 312–334 (2001).
- Buijs, R. M. The anatomical basis for the expression of circadian rhythms: the efferent projections of the suprachiasmatic nucleus. *Prog. Brain Res.* **111**, 229–240 (1996).
- Albrecht, U., Sun, Z. S., Eichele, G. & Lee, C. C. A differential response of two putative mammalian circadian regulators, *mper1* and *mper2*, to light. *Cell* **91**, 1055–1064 (1997).
- Tei, H. *et al.* Circadian oscillation of a mammalian homolog of the *Drosophila period* gene. *Nature* **389**, 512–516 (1997).
- Miyamoto, Y. & Sancar, A. Vitamin B2-based blue-light photoreceptors in the retinohypothalamic tract as the photoactive pigments for setting the circadian clock in mammals. *Proc. Natl Acad. Sci. USA* **95**, 6097–6102 (1998).
- Zylka, M. J., Shearman, L. P., Weaver, D. R. & Reppert, S. M. Three *period* homologs in mammals: differential light responses in the suprachiasmatic circadian clock and oscillating transcripts outside of brain. *Neuron* **20**, 1103–1110 (1998).
- LeSauter, J. & Silver, R. Output signals of the SCN. *Chronobiol. Int.* **15**, 535–550 (1998).
- Winer-Serhan, U. H., Broide, R. S., Chen, Y. & Leslie, F. M. Highly sensitive radioactive *in situ* hybridization using full length hydrolyzed riboprobes to detect alpha 2 adrenoceptor subtype mRNAs in adult and developing rat brain. *Brain Res. Brain Res. Protocols* **3**, 229–241 (1999).

Supplementary Information accompanies the paper on *Nature's* website (<http://www.nature.com>).

Acknowledgements

We thank Y. Chen, D. Lin, H. Nagasaki and L. Shearman for technical assistance; S. Loughlin, S. Reppert and P. Sassone-Corsi for discussions; and B. Semler for access to a luminometer. We also thank J. Takahashi, M. Vitaterna and B. van der Horst for providing access to mutant mice. The research is partially supported by grants from the National Institutes of Health. C.M.B. is a recipient of a UC Irvine MSTP training grant.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to Q.-Y.Z. (e-mail: qzhou@uci.edu).

Adaptation in the chemotactic guidance of nerve growth cones

Guo-li Ming^{*†}, Scott T. Wong^{‡‡}, John Henley^{†‡}, Xiao-bing Yuan^{‡§}, Hong-jun Song^{||}, Nicholas C. Spitzer^{*} & Mu-ming Poo[‡]

^{*} Section of Neurobiology, Division of Biology, University of California at San Diego, La Jolla, California 92093, USA

[‡] Division of Neurobiology, Department of Molecular and Cell Biology, University of California, Berkeley, California 94720, USA

^{||} The Salk Institute for Biological Studies, La Jolla, California 92037, USA

[†] These authors contributed equally to this work

[§] Present address: Institute of Neuroscience, Shanghai Institute of Biological Sciences, Chinese Academy of Sciences, Shanghai 100032, China

Pathfinding by growing axons in the developing nervous system may be guided by gradients of extracellular guidance factors. Analogous to the process of chemotaxis in microorganisms, we found that axonal growth cones of cultured *Xenopus* spinal neurons exhibit adaptation during chemotactic migration, undergoing consecutive phases of desensitization and resensitization in the presence of increasing basal concentrations of the guidance factor netrin-1 or brain-derived neurotrophic factor. The desensitization is specific to the guidance factor and is accompanied by a reduction of Ca^{2+} signalling, whereas resensitization requires activation of mitogen-associated protein kinase and local protein synthesis. Such adaptive behaviour allows the growth cone to re-adjust its sensitivity over a wide range of concentrations of the guidance factor, an essential feature for long-range chemotaxis.

The formation of specific connections between nerve cells in the developing nervous system depends on the ability of growing axons to reach appropriate target cells, a process known as pathfinding^{1,2}. Intrigued by the amoeboid morphology of axonal growth cones, Ramón y Cajal suggested more than a century ago that growth cones might be guided by gradients of chemical substances secreted from the target tissue in a manner similar to the chemotaxis of leukocytes³. There is now substantial evidence that axon pathfinding in many regions of the developing brain is indeed guided by diffusible guidance cues^{1,2,4–7}. For long-range guidance based on the detection of chemoattractant gradients, the growth cone must be able to respond reliably to small gradients of guidance cues across its surface. This may be achieved by amplification of guidance signals through cellular transduction processes⁸. In addition, as the growth cone migrates in an environment in which the basal concentration of the guidance cue varies by many orders of magnitude, it may also need to constantly re-adjust its sensitivity, a process commonly referred to as adaptation^{8–11}. The cellular mechanisms underlying the amplification and adaptation of guidance signals are largely unknown. In this study, we have examined quantitatively the phenomenon of adaptation in the chemotaxis of nerve growth cones *in vitro*. When exposed to high levels of netrin-1 or brain-derived neurotrophic factor (BDNF), the growth cones of cultured *Xenopus* spinal neurons undergo consecutive phases of desensitization (or interference) and resensitization (or adaptation) in their ability to detect gradients of netrin-1 or BDNF. This process is reminiscent of the adaptation in bacterial chemotaxis towards a chemoattractant¹². These cyclic phases of desensitization and resensitization are also reflected in the 'zig-zag' path of alternating attractive and repulsive turning of fast-advancing growth cones in their migration towards the source of these two guidance cues (netrin-1 and BDNF). Further studies indicate that desensitization occurs at a level upstream from the signalling effector(s) shared by netrin-1 and BDNF, and is accompanied by a loss of Ca^{2+} signalling in response to the gradient of the guidance factor. The resensitization, on the other hand, requires the activation of mitogen-associated protein kinase (MAPK), which is induced by both netrin-1 and BDNF. Surprisingly, we found that inhibition of protein translation, but not gene transcription, blocked the resensitization of the growth cone to the netrin-1 gradient, suggesting that MAPK-dependent initiation of protein synthesis is required for the resensitization

process. This idea is further supported by the observation that inhibition of protein synthesis in a growing neurite severed from the cell body abolished the chemotaxis of its growth cone in a netrin-1 gradient.

Adaptation in growth-cone chemotaxis

Growth cones of isolated *Xenopus* spinal neurons in culture exhibit chemotactic turning responses when exposed to a gradient of netrin-1 (ref. 13)—a diffusible guidance cue for developing axons⁴ and migrating neuroblasts¹⁴. The gradient was produced by repetitive pulsatile ejection of picolitres of solution containing netrin-1 ($5 \mu\text{g ml}^{-1}$) from a micropipette^{15–18}, which was positioned at a 45° angle with respect to the direction of initial neurite extension and a distance of $100 \mu\text{m}$ away from the centre of the growth cone. As shown in Fig. 1a, marked chemotactic turning of the growth cone was observed 30 min after application of this standard netrin-1 gradient. To examine the role of the basal level of receptor activation on the turning response of the growth cone, we uniformly applied netrin-1 at different concentrations to the culture immediately before the onset of the turning assay. We found that the overall chemotactic response, as determined by the average turning angle of growth cones at the end of the 30-min assay, was significantly reduced in the presence of 0.5 ng ml^{-1} netrin-1 in a bath solution, and completely abolished when the bath concentration was increased to 1 or 5 ng ml^{-1} (Figs 1b, d and 2a). The presence of netrin-1 in the bath interfered and desensitized the ability of the growth cone to sense the standard netrin-1 gradient, probably owing to uniform activation of netrin-1 receptors before exposure to the gradient. This dose-dependent desensitization of chemotactic turning was also observed for BDNF, a neurotrophin known to be a chemoattractant for some growth cones *in vitro*^{17,18} and *in vivo*¹⁹. As shown in Fig. 2c, BDNF at a bath concentration of 20 ng ml^{-1} completely desensitized the attractive response induced by the standard gradient of BDNF ($50 \mu\text{g ml}^{-1}$ in the pipette). The desensitization induced by either netrin-1 or BDNF required the continuous presence of the factor. Removal of netrin-1 from the medium after 30 min pre-incubation resulted in the recovery of the turning response within 30 min (Fig. 2c), the minimal time required for the turning assay. At the threshold bath concentration of netrin-1 (1 ng ml^{-1}), the lack of response towards the standard netrin-1 gradient represents a reduction rather than complete loss of netrin-

1 sensitivity, as increasing the gradient of netrin-1 by putting the netrin-1 pipette at 70 μm from the growth cone resulted in a significant turning response in the presence of 1 ng ml^{-1} of netrin-1 (Fig. 2a).

In the continuous presence of netrin-1 applied to the bath (at 1 or 5 ng ml^{-1}), which desensitized the ability of the growth cone to detect the standard netrin-1 gradient, the growth cone regained its sensitivity to the gradient after a prolonged period of incubation. When we applied a netrin-1 gradient at 90 or 120 min after bath-application of netrin-1 (1 or 5 ng ml^{-1}), robust attractive turning of the growth cone was found (Figs 1c, d and 2c). This recovery of sensitivity was not due to degradation of netrin-1 in the bath, as the medium containing 1 ng ml^{-1} netrin-1 remained fully effective in desensitizing the growth cone when introduced into fresh cultures (see Methods). Furthermore, when the growth cone recovered its sensitivity towards the standard netrin-1 gradient after 90 min incubation of 1 ng ml^{-1} in the bath, increasing netrin-1 in the bath to 5 ng ml^{-1} again desensitized the growth cone (Fig. 2c). Attractive turning in the standard gradient of BDNF also recovered

in the continuous presence of 20 ng ml^{-1} BDNF, with normal attractive turning observed when the turning assay began at 60 min after application of BDNF to the bath (Fig. 2d). Thus the growth cone is capable of re-adjusting its sensitivity towards a guidance factor after a change in the basal concentration of the factor.

Zig-zag chemotaxis of the growth cone

Does the adaptive process revealed by bath application of guidance cues have a role in growth-cone chemotaxis in a gradient of the guidance cue? To address this question, we further examined the behaviour of the growth cone as it migrates up the gradient of netrin-1 or BDNF. As shown in Fig. 3, we observed that fast-advancing growth cones consistently exhibited a zig-zag pattern of migration towards the source of netrin-1 or BDNF, with alternating periods of attractive and repulsive responses. Whereas the behaviour of filopodia initiation and the path of growth-cone migration showed a zig-zag pattern, the trajectory of the neurite became smoother with time. Thus the zig-zag pattern was often not detectable for slowly advancing neurites at the end of the turning assay. The average interval for the zig-zag was 20.0 ± 2.5 min (s.e.m., $n = 5$) in the standard netrin-1 gradient and 22.8 ± 2.0 min (s.e.m., $n = 5$) in the standard BDNF gradient. In the absence of the guidance factor, there is no regular zig-zag and the growth cone advances in an irregular, unpredictable manner. This irregular path is probably determined by variations in substrate adhesion and other random variables in the culture condition. As discussed below, this zig-zag pattern of extension is consistent with cycles of rapid and dose-dependent desensitization and slower resensitization/adaptation processes at the growth cone. Asymmetry in the extent of desensitization across the growth cone, with a higher level of desensitization on the side facing the gradient may be the cause of repulsion after each attractive turning. The average

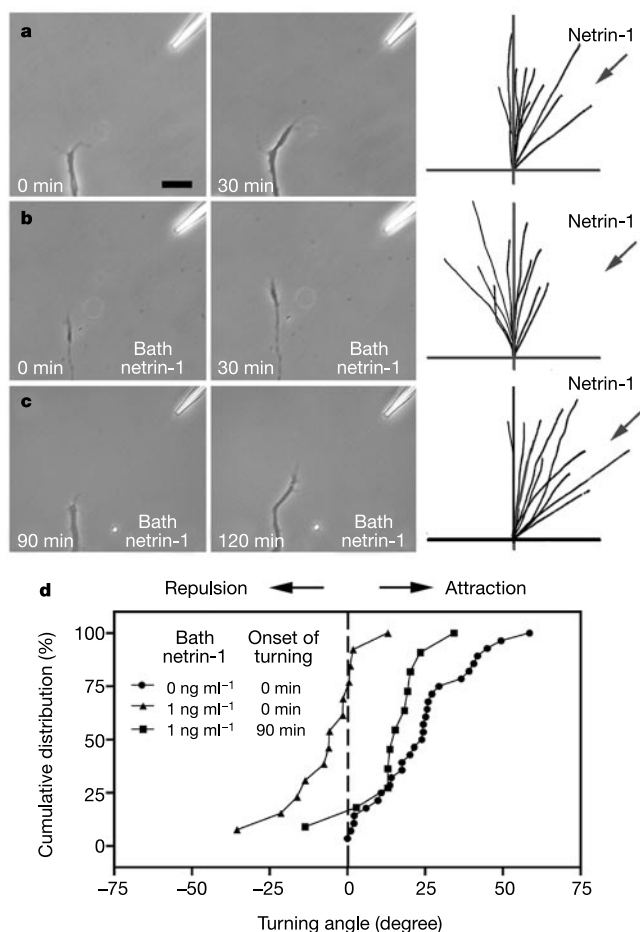


Figure 1 Turning of nerve growth cones induced by a gradient of netrin-1. **a**, Images of a growth cone at the onset (left) and the end (right) of a 30-min exposure to the standard netrin-1 gradient (arrow). Superimposed traces depict ten samples of neurite trajectories at the end of the assay (diagram). The origin is the centre of the growth cone at the onset of the assay, and the original direction of growth was vertical. Scale bar, 20 μm for images and 5 μm for diagram. **b**, **c**, As in **a**, except that netrin-1 (1 ng ml^{-1}) was added to the culture immediately before (**b**) or 90 min before (**c**) the onset of the netrin-1 gradient. **d**, The distribution of turning angles. For each experimental condition, angular positions of all growth cones at the end of 30-min exposure to a netrin-1 gradient are plotted. The per cent value refers to the percentage of growth cones with turning angle less than or equal to a given angular value. Data are from experiments similar to those shown in **a**–**c**.

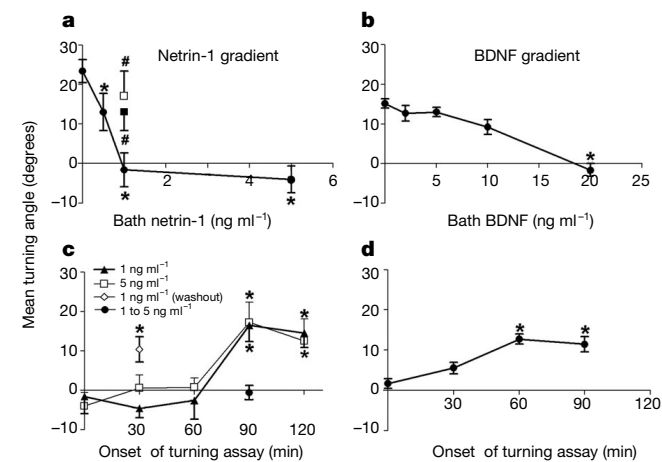


Figure 2 Desensitization and resensitization of turning responses. **a**, **b**, Dose-dependent desensitization of growth-cone responses in a standard gradient of netrin-1 (**a**) or BDNF (**b**). The concentration refers to netrin-1 and BDNF added in the saline before the turning assay. The data shown are mean turning angles ($\pm$ s.e.m., $n = 10$ –30). An asterisk indicates values significantly different from controls (without netrin-1 or BDNF in the bath; $P < 0.01$). Also shown is turning induced by a higher netrin-1 gradient (pipette at 70 μm) in the absence (open square) or presence (filled square) of bath netrin-1 (1 ng ml^{-1}). Hash symbols indicate values significantly different from controls (standard netrin-1 gradient with 1 ng ml^{-1} bath netrin-1). **c**, **d**, Time course of resensitization. The mean turning angle was measured at various times ($n = 10$ –30) after addition of netrin-1 (1 and 5 ng ml^{-1} , in **c**) or BDNF (20 ng ml^{-1} , in **d**) to the bath. An asterisk indicates values significantly different from the control ($t = 0$ min, $P < 0.01$). Washout, bath netrin-1 (1 ng ml^{-1}) was washed out after 30 min; 1 to 5, bath netrin-1 was changed from 1 to 5 ng ml^{-1} after 90 min incubation before the turning assay.

interval between each zig-zag was shorter than the time required for resensitization described above. Owing to the shallow gradient (pipette at 100 μm) used in the standard turning assay, the time course of resensitization found in the bath incubation experiments may represent an underestimate of the normal adaptive responses of the growth cone when it faces progressively steeper gradients during its migration towards the pipette.

Heterologous desensitization

The specificity of growth-cone adaptation was further studied by examining heterologous cross-desensitization between netrin-1 and BDNF. For transduction of the directional signal triggered by a gradient of receptor activation at the growth cone, a gradient of activity of cytoplasmic effector(s) must be generated⁸. As these two guidance cues share common cytosolic transduction mechanisms, for example, activation of phosphoinositide 3 kinase (PI(3)K) and phospholipase C γ (PLC- γ , ref. 20), and also Ca^{2+} signalling^{13,17,18,21}, uniform bath application of one factor at concentrations that cause saturated activation of the shared effector(s) should desensitize the response of the growth cone towards a gradient of the other factor. In agreement with a previous report²⁰, we found that the growth cone exhibited no turning response towards the standard netrin-1 gradient when 100 ng ml^{-1} BDNF was present in the bath (Fig. 4a). Conversely, a high bath concentration of netrin-1 (10 ng ml^{-1}) desensitized the turning response of the growth cone towards the standard BDNF gradient (Fig. 4b). Thus, saturated activation of guidance signalling mechanisms indeed results in heterologous desensitization of the growth cone. However, when the bath concentration of BDNF was lowered to the threshold level (20 ng ml^{-1}) for inducing homologous desensitization towards the standard BDNF gradient, we found significant turning response still occurred towards the standard netrin-1 gradient, although

the average turning angle was lower than that found in the absence of BDNF in the bath (Fig. 4a). Conversely, bath application of 1 ng ml^{-1} netrin-1 resulted in no heterologous desensitization of the growth cone towards the BDNF gradient (Fig. 4b). Thus homologous desensitization induced by bath application of netrin-1 or BDNF at the threshold concentration did not cause saturated activation of their respective surface receptors nor their shared cytoplasmic effector(s). This is consistent with the finding that after homologous desensitization of the growth cone towards the standard netrin-1 gradient at the threshold bath concentration of netrin-1 (1 ng ml^{-1}), increasing the slope of the netrin-1 gradient resulted in significant attractive turning (Fig. 2a). The absence of heterologous desensitization described above also suggests that homologous desensitization of netrin-1 and BDNF responses at threshold concentrations is specific to the factor and occurs upstream from the shared cytoplasmic effector(s), possibly at the level of membrane receptors.

Cytosolic Ca^{2+} signalling

Cytosolic Ca^{2+} ($[\text{Ca}^{2+}]_i$) regulates neurite outgrowth^{22–25} and mediates growth-cone turning responses induced by acetylcholine¹⁶ and netrin-1 (refs 13, 21). We monitored changes of growth cone $[\text{Ca}^{2+}]_i$ induced by the standard netrin-1 gradient, using the Ca^{2+} -sensitive fluorescence indicator Oregon green BAPTA-1 conjugated to dextran (Fig. 5a, b). Elevation of $[\text{Ca}^{2+}]_i$ at the growth cone was consistently observed within minutes after the onset of the gradient, with a transient gradient of Ca^{2+} across the growth cone detected in some cases²¹. Moreover, we found that bath application of netrin-1 at a desensitizing concentration (5 ng ml^{-1}) caused a persistent $[\text{Ca}^{2+}]_i$ elevation (Fig. 5d) and the standard netrin-1 gradient failed to induce significant $[\text{Ca}^{2+}]_i$ elevation during the period of desensitization (Fig. 5a, c, e). This elevation of $[\text{Ca}^{2+}]_i$ induced by bath-applied netrin-1 was not responsible for triggering the desensitization. We found that similar $[\text{Ca}^{2+}]_i$ elevation induced by perfusing the culture with medium containing 2.5 mM Ca^{2+} and

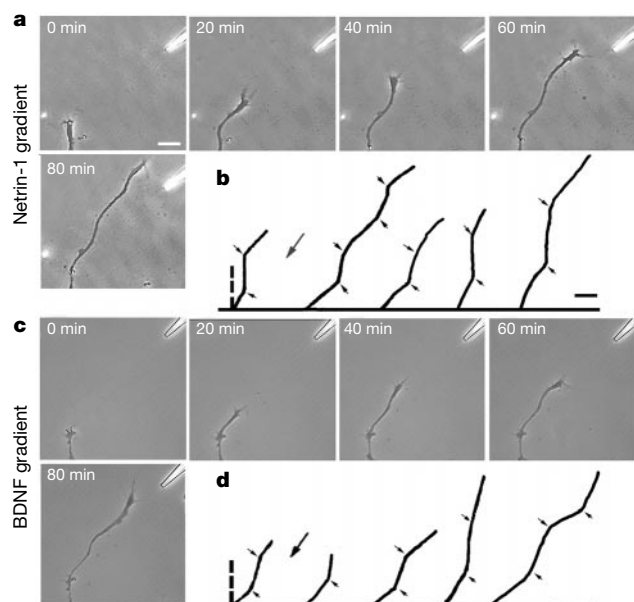


Figure 3 Zig-zag pathfinding of the growth cone. **a, b**, Images of fast-advancing growth cones in a standard gradient of netrin-1 at different times after the onset of the gradient (**a**). Note the zig-zag pattern of the neurite trajectory, reflecting alternating attractive and repulsive turning of the growth cone. Sample tracings of the migratory path of growth cones examined in similar fashion are indicated (**b**). The original direction of growth was vertical and the large arrow marks the direction of the gradient (shown only for one growth cone). Note that the tracing depicts the actual position of five different growth cones during chemotaxis, not the neurite trajectory at the end of the experiment. Small arrows indicate turning points of the zig-zag. Scale bar, 20 μm for images and 5 μm for traces. **c, d**, Similar to **a, b**, except that a standard BDNF gradient was used.

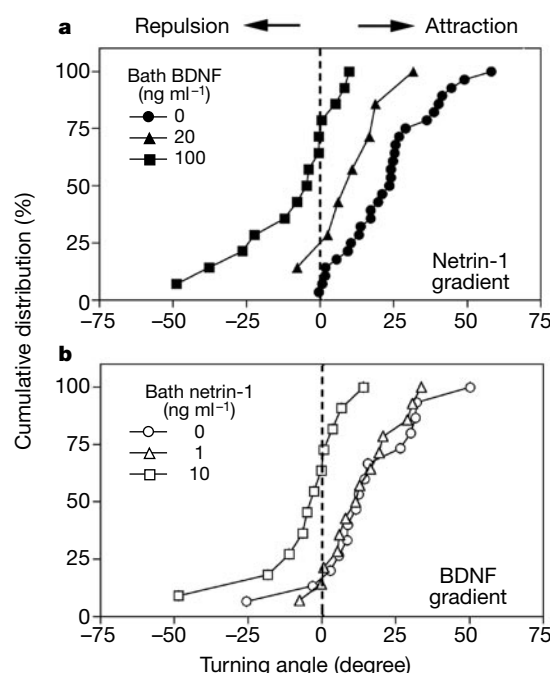


Figure 4 Heterologous desensitization between netrin-1 and BDNF. **a, b**, Growth-cone turning induced by the standard gradient of netrin-1 or BDNF, in the presence of different concentrations of BDNF or netrin-1 in the bath, respectively. Note significant heterologous desensitization at the high but not at low concentrations of BDNF and netrin-1 ($P < 0.01$).

12.5 mM K^+ had no significant effect on the attractive turning of the growth cone in the netrin-1 gradient (data not shown). Thus, the growth-cone sensitivity to the netrin-1 gradient correlates with Ca^{2+} signalling, but $[Ca^{2+}]_i$ elevation by itself is not sufficient to trigger desensitization of the growth cone.

After 90 min incubation with netrin-1 in the bath, when resensitization of the growth cone had occurred, we found that the standard netrin-1 gradient was able to induce further $[Ca^{2+}]_i$ elevation, which was superimposed on top of the elevated basal $[Ca^{2+}]_i$ (Fig. 5b, c, e). Similar results were obtained with ratio imaging of fluorescence from co-injected Oregon green and Texas red (see Methods), making it unlikely that these results were due to fluctuations in the growth-cone thickness or changes in focal plane that may have occurred during the experiment. Taken together, these findings indicate that Ca^{2+} signals correlate with

the sensitivity of the growth cone, and that netrin-1-induced desensitization/resensitization of the turning responses results from transduction failure/restoration at the level(s) upstream from Ca^{2+} signalling.

Resensitization requires MAPK activation

To further explore the intracellular events involved in the resensitization of chemotactic responses, we examined the role of MAPK signalling, which is known to be activated by neurotrophins²⁶. We treated *Xenopus* spinal cord explant cultures with BDNF (10 ng ml^{-1}) for various times and examined MAPK activity by monitoring phosphotransferase activity of immunoprecipitated MAPK towards myelin basic protein, a standard substrate for examining MAPK activity²⁷. We found that BDNF induced a robust increase in MAPK activity that persisted for at least 60–90 min (Fig. 6). Interestingly, MAPK activity was also activated by netrin-1 (20 ng ml^{-1}) by 4–5-fold, peaking within 5–10 min after the onset of the treatment. Similar results were obtained by immunoblot analysis using a specific antibody directed against dually phosphorylated MAPK (Fig. 6c). Furthermore, the MAPK activation was blocked by U0126 ($10 \mu\text{M}$), which is a potent and selective inhibitor of MEK1, the upstream activator of MAPK²⁸.

Is MAPK activation by netrin-1 required for resensitization of the turning response? We found that bath application of $10 \mu\text{M}$ U0126 or PD98059 (ref. 29)—two structurally different inhibitors for MEK1—for the first 60 min completely abolished resensitization of the turning response when tested at 90 min after bath application of netrin-1 (1 ng ml^{-1} , Fig. 6d). Furthermore, in the absence of any bath-applied netrin-1, the attractive turning response induced by the standard netrin-1 gradient was significantly attenuated when the turning assay was performed 20 min after addition of these MEK1 inhibitors into the culture (Fig. 6e). Pre-incubation with U0126 for 60 min before the turning assay, however, completely abolished the chemotaxis in the netrin-1 gradient (Fig. 6e). These results suggest that the effects of MAPK inhibition may proceed with a slow time course, resulting in a lack of resensitization required for effective chemotaxis of the growth cone.

The role of local protein synthesis

The MAPK pathway has been implicated in a wide range of cellular processes, including stimulation of protein synthesis in response to growth factors³⁰. We thus examined whether resensitization of the growth cone response to gradients of netrin-1 requires protein synthesis. Neurons were incubated for 60 min with netrin-1 (1 ng ml^{-1}) and cycloheximide ($10 \mu\text{g ml}^{-1}$), a reversible protein synthesis inhibitor³¹. They were then perfused with fresh saline containing 1 ng ml^{-1} netrin-1 for 30 min and assayed for the turning response at 90 min, a time normally sufficient for resensitization. We found that the growth cones exhibited no recovery in their sensitivity towards the standard netrin-1 gradient (Fig. 7a). Similar results were also observed with anisomycin ($30 \mu\text{M}$, Fig. 7a), another reversible protein synthesis inhibitor³². In contrast, when 5,6-dichlorobenzimidazole riboside (DRB, $50 \mu\text{M}$), a reversible inhibitor of gene transcription³³, was used instead of cycloheximide or anisomycin in the incubation saline, we observed normal recovery of chemoattraction when assayed at 90 min (Fig. 7a). Importantly, the control experiments using cultures not incubated with netrin-1 in the bath showed that such a 60-min treatment with cycloheximide or DRB did not significantly affect the attractive turning response induced by the netrin-1 gradient examined 30 min after washing away the drug (Fig. 7a). Thus, protein synthesis but not gene transcription during the first 60 min after the treatment of a desensitizing concentration of netrin-1 is required for resensitization to take place.

The role of protein synthesis was further examined for chemotaxis of the growth cone under the standard netrin-1 gradient in the absence of any bath-applied netrin-1. As shown in Fig. 7b, after

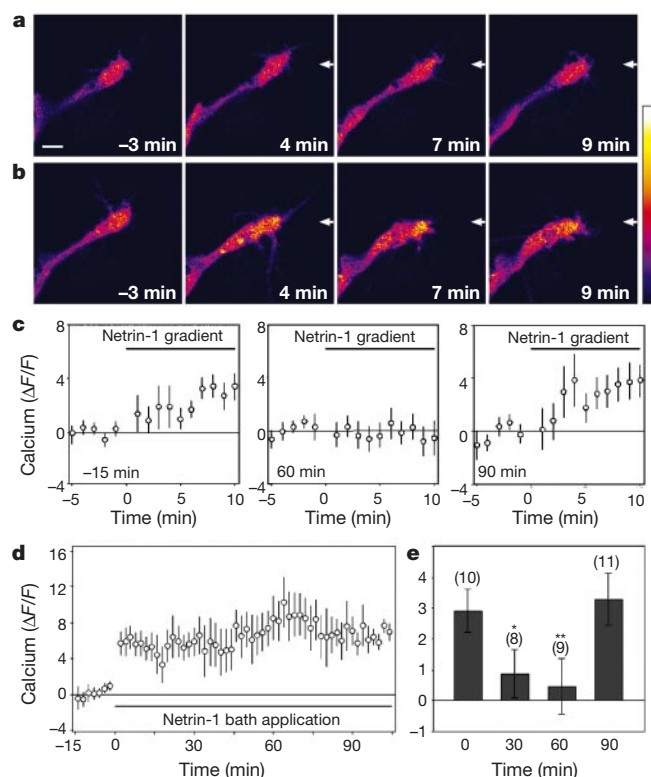


Figure 5 Calcium signals induced by a netrin-1 gradient correlate with adaptation. **a, b**, Fluorescence images of a *Xenopus* growth cone loaded with a Ca^{2+} -sensitive fluorescent indicator (Oregon green BAPTA-1-dextran; images show Oregon green BAPTA fluorescence normalized to Texas red). The neuron was incubated with netrin-1 (5 ng ml^{-1}) in the bath 30 min (**a**) or 90 min (**b**) before exposure to a netrin-1 gradient (arrows). Images show Oregon green fluorescence at different times (min) before and after the onset of the gradient. Pseudocolour indicates Ca^{2+} levels: white is highest and black is lowest. Scale bar, $10 \mu\text{m}$. **c**, Percentage changes in fluorescence ($\Delta F/F$) with time of the growth cone induced by the netrin-1 gradient (bar) 15 min before (left), and 60 min (middle) and 90 min (right) after addition of netrin-1 (5 ng ml^{-1}) to the bath. The fluorescence values (1-min bins) were normalized to the mean value during the 5-min period before applying the gradient for each growth cone. Data points represent mean $\pm$ s.e.m. ($n = 9$ –11). **d**, Ca^{2+} changes in the growth cone with time before and after uniform exposure to 5 ng ml^{-1} netrin-1 in the bath solution. **e**, Summary of Ca^{2+} changes induced by a netrin-1 gradient after pre-incubation with netrin-1 (5 ng ml^{-1}) in the bath for 0, 30, 60 and 90 min. The mean fluorescence change in the growth cone measured during the last 5 min of a 10-min exposure to a netrin-1 gradient was normalized to the mean pre-gradient value over a 10-min period. Data significantly different from the control (no netrin-1 in the bath) are marked by a single ($P < 0.05$) or double ($P < 0.01$) asterisk, using the Mann–Whitney rank sum test. Values in parentheses indicate total number of growth cones examined in this condition.

20 min pre-incubation of cycloheximide ($10 \mu\text{g ml}^{-1}$), the growth cone consistently failed to exhibit chemoattraction in the standard netrin-1 gradient; instead, repulsive turning was observed in most growth cones. Similar results were found when cultures were treated with $30 \mu\text{M}$ anisomycin or $100 \mu\text{M}$ emetine³⁴, two other specific protein synthesis inhibitors. On the other hand, treatment with 10 ng ml^{-1} actinomycin³⁵, or $50 \mu\text{M}$ DRB, two specific transcriptional inhibitors, had no effect on chemoattraction induced by netrin-1 (Fig. 7b). These observations are consistent with the idea that in order for continuous chemotaxis to be maintained, protein synthesis is necessary for resensitization of the growth cone towards netrin-1. As discussed below (see Fig. 8), we attribute the repulsive turning to an asymmetric desensitization across the growth cone in the netrin-1 gradient. In the absence of resensitization, the growth cone may extend in the direction that was set after the first cycle of desensitization, leading to an overall repulsive response.

Protein synthesis responsible for the resensitization of the growth cone may be carried out locally in the neurite. To test this hypothesis, we transected the neurite from the neuronal soma and examined the turning response of the growth cone in a gradient of netrin-1. Out of all isolated neurites, a fraction of them (18 out of 75) recovered their growth-cone motility in culture. These isolated growing neurites responded to the gradient of netrin-1 with attractive turning (Fig. 7c). However, when cycloheximide ($10 \mu\text{g ml}^{-1}$) was added to the bath after the transection, these

growth cones failed to exhibit chemotaxis in the netrin-1 gradient and instead showed repulsion (Fig. 7d). This repulsion is consistent with a lack of resensitization or adaptation after the initial asymmetric desensitization, causing a persistent state of repulsion at the end of the first zig-zag (see Fig. 8). Thus, local protein synthesis in the neurite appears to be responsible for resensitization of the growth cone towards netrin-1, a process that is required for effective long-range chemotaxis.

Discussion

Adaptive behaviour in growth-cone guidance has been observed in retinal ganglion axons, which can adjust their sensitivity according to the basal concentration of the tectal membrane repellents so that they grow up the gradient of repellents for a fixed increment of concentration before stalling^{10,11}. Adaptation in both chemotaxis and chemorepulsion requires a modulation in the gain of guidance-signal transduction⁸. In analogy to the desensitization of β -adrenergic receptors³⁶, the guidance cue may induce a modification of its receptors or receptor-interacting proteins, resulting in changes in the ligand affinity or the efficacy of signal transduction. In response to increasing basal concentrations of netrin-1 or BDNF, the receptors may also be downregulated in a manner similar to that found for many growth factors³⁷. Subsequent recovery of functional receptors, in a MAPK- and protein synthesis-dependent manner³⁸, may lead to resensitization. In addition to changes at the receptor

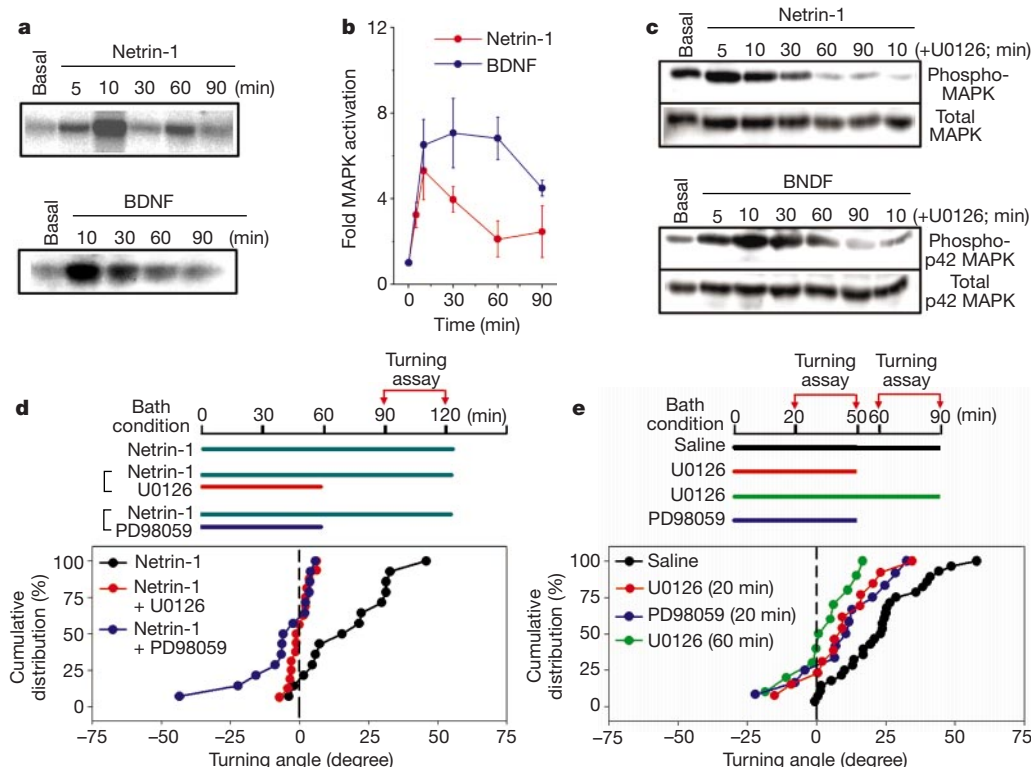


Figure 6 Activation of MAPK by netrin-1 and BDNF, and involvement of MAPK in growth-cone chemotaxis. **a**, Time course of netrin-1- and BDNF-stimulated MAPK activity towards myelin basic protein. Activities at given time points were visualized by separation on 10% SDS-polyacrylamide gel electrophoresis gels followed by autoradiography. **b**, Radioactivity from excised gel bands as shown in **a** was quantified by dividing the increase in c.p.m. (counts per minute) at each time point by the basal c.p.m. levels. All data are means $\pm$ s.e.m. from at least three experiments. **c**, MAPK activation in response to netrin-1 and BDNF was confirmed by immunoblot analysis with phospho-specific antibodies directed against the active form of MAPK at different time points as indicated. U0126 ($10 \mu\text{M}$), a specific MEK1 inhibitor, blocked MAPK activation. **d**, Effects of MAPK inhibitor on resensitization/adaptation of the growth cone. U0126 ($10 \mu\text{M}$) or PD98059

($10 \mu\text{M}$) was applied during the first 60 min of bath incubation with netrin-1 (1 ng ml^{-1}), and the turning assay was performed at $t = 90\text{--}120$ min. Data shown are distributions of turning angles for growth cones induced by the standard gradient of netrin-1 under different conditions. The distributions of turning angle in the presence of inhibitors were significantly different from that observed in the absence of the drugs ($P < 0.01$). **e**, Effects of MAPK inhibitors on the netrin-1-induced turning response. Shown are the distributions of turning angles for growth cones induced by the standard netrin-1 gradient in the absence or presence of MAPK inhibitor U0126 ($10 \mu\text{M}$) or PD98059 ($10 \mu\text{M}$). These inhibitors were present 20 or 60 min before and during the turning assay. The distributions of turning angle in the presence of inhibitors were significantly different from that observed in the absence of the inhibitors ($P < 0.01$).

level, modulation of downstream cytoplasmic effectors may alter the sensitivity of the growth cone to the guidance cue^{8,39}. For example, elevation of cytosolic cyclic AMP can boost the chemotactic response of the growth cone in a gradient of low concentration of netrin-1 (refs 13, 40) or BDNF¹⁷. The absence of heterologous desensitization at a threshold bath concentration of the factor that is sufficient for homologous desensitization suggests that homologous desensitization may occur upstream from cytoplasmic effector(s) shared by these two guidance cues, an idea further supported by results from our Ca^{2+} imaging studies.

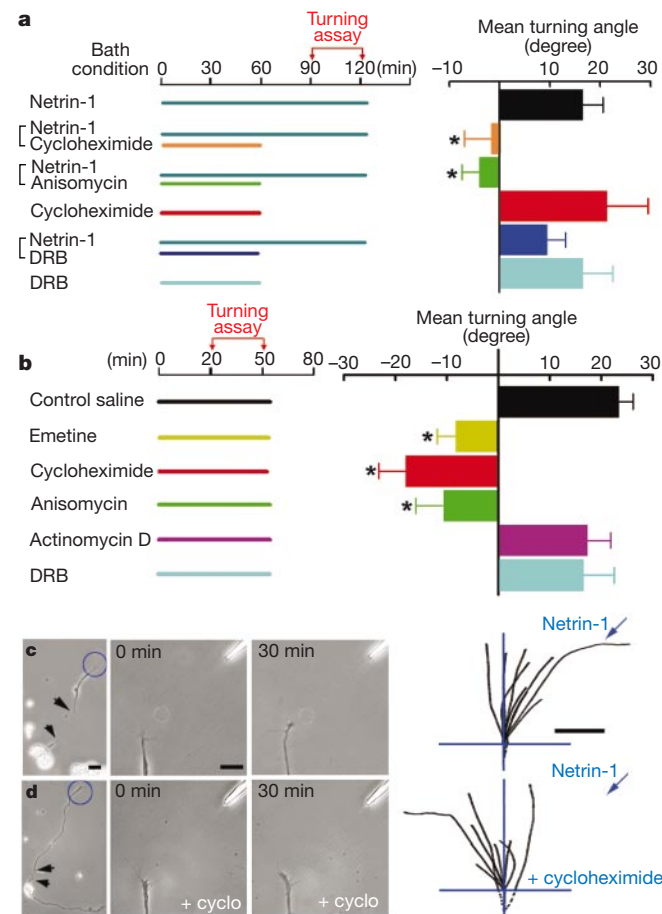


Figure 7 Effects of protein synthesis inhibition. **a**, Reversible inhibitors of protein synthesis (cycloheximide, $10 \mu\text{g ml}^{-1}$ or anisomycin, $30 \mu\text{M}$) or gene transcription (DRB, $50 \mu\text{M}$) were applied during the first 60 min of incubation with netrin-1 (1 ng ml^{-1} in the bath), and the mean turning angle ($\pm$ s.e.m., $n = 8-15$) was determined by the standard netrin-1 gradient at $t = 90-120$ min. Data are also shown for turning assayed at 90 min with bath application of netrin-1 (1 ng ml^{-1} in the bath) alone, or after 60-min treatment with the inhibitor in the absence of bath netrin-1. An asterisk indicates values significantly different from the control with a bath application of netrin-1 alone ($P < 0.01$). **b**, Mean turning angles ($\pm$ s.e.m., $n = 8$ to 15) of growth cones in the standard netrin-1 gradient in the absence or presence of protein synthesis inhibitors emetine ($100 \mu\text{M}$), cycloheximide ($10 \mu\text{g ml}^{-1}$) and anisomycin ($30 \mu\text{M}$), or transcription inhibitors actinomycin D (10 ng ml^{-1}) and DRB ($50 \mu\text{M}$). Each drug was present for 20 min before and throughout the turning assay. An asterisk indicates values significantly different from the control group (no drugs in the bath; $P < 0.01$). **c**, Neurites exhibit normal chemotaxis in the standard netrin-1 gradient, after being mechanically transected near the cell body (arrowheads). The growth cone of the severed neurite (blue circle) was exposed to the netrin-1 gradient; images are shown at 0 and 30 min after the onset of the gradient. Scale bar, $20 \mu\text{m}$. The superimposed traces show neurite trajectories of all isolated neurites that resumed growth ($n = 9$; scale bar, $5 \mu\text{m}$). **d**, Similar to **c**, except that cycloheximide ($10 \mu\text{g ml}^{-1}$) was added after neurite transection, at 20 min before and during the turning assay.

The resensitization of growth cones towards netrin-1 and BDNF was inhibited by two structurally different MEK1 inhibitors, suggesting that MAPK activation is critical for resensitization. Activation of MAPK seems to provide either an enhanced recruitment or a *de novo* synthesis of chief components required for resensitization or adaptation. Although it is well known that MAPKs are direct downstream effectors of BDNF TrkB signalling²⁶, how netrin-1 activates MAPK is unclear. The MAPK-dependent adaptation towards chemoattractants seems to be a phylogenetically conserved mechanism. In amoeboid chemotaxis in *Dictyostelium*, MAPK is also required for adaptation of the cells to high concentrations of the chemoattractant cAMP, but not for the process of locomotion itself⁴¹. The effect of MAPK inhibition on growth-cone chemotaxis is a slow process (Fig. 6e), consistent with a blockade of resensitization, although some effects or immediate turning response may also be present. Our findings on the effect of translation inhibitors suggest that constitutive protein synthesis in the neurite is critical for resensitization. As MAPK activity can initiate protein translation⁴², and netrin-1 and BDNF can induce MAPK activation (Fig. 6) and trigger protein synthesis in cell cultures^{43,44}, it is likely that MAPK-dependent local protein synthesis in the neurite⁴⁵ is responsible for the resensitization process.

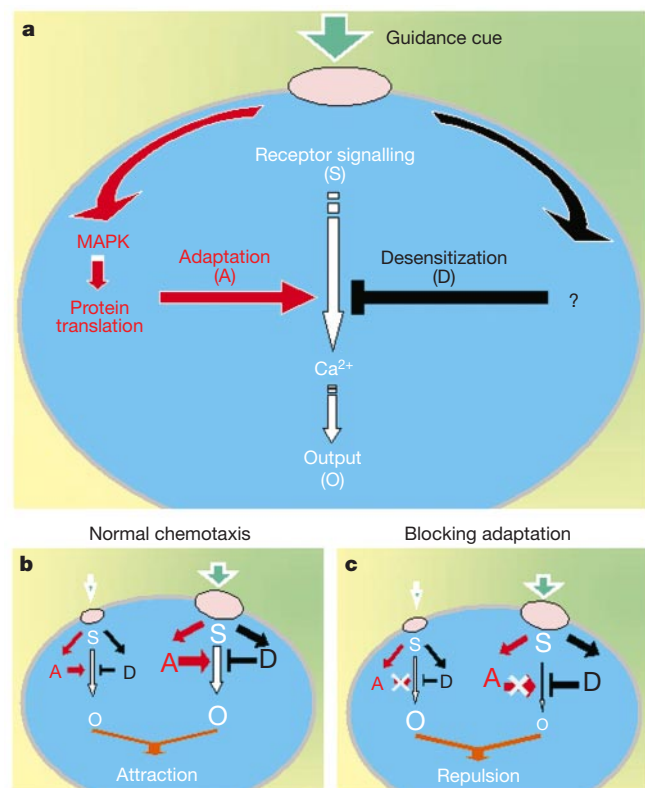


Figure 8 A model of growth-cone adaptation in a gradient of guidance cues. **a**, The guidance cue triggers a cytoplasmic signal (S) and two parallel processes of desensitization (D) and resensitization or adaptation (A), which regulate the output (O) of the signalling cascade, mediated through elevation of Ca^{2+} , leading to cytoskeletal rearrangements associated with attractive or repulsive responses. **b**, Asymmetry in the output determines the direction of growth. With time, the effect of resensitization or adaptation compensates for that of desensitization, leading to chemotaxis. **c**, Assuming the onset of desensitization precedes that of resensitization or adaptation, the asymmetric desensitization (higher towards the gradient) before the onset of resensitization or adaptation results in repulsive turning after the initial attraction, hence the zig-zag trajectory of growth (Fig. 3). Blocking resensitization or adaptation by protein synthesis inhibitors results in a persistent asymmetric output after the initial desensitization, leading to the end result of repulsion (Fig. 7b, d).

To account for the zig-zag pattern of growth-cone migration and the repulsive turning in the presence of translation inhibitors, we propose that, in the presence of a netrin-1 gradient, the difference in receptor activation on two sides of the growth cone leads to an asymmetry in the receptor activation for cytoplasmic signalling and in the rate or the extent of desensitization (Fig. 8). Owing to differences in the time course of desensitization and resensitization, such asymmetry in desensitization in fast-advancing growth cones may result in a repulsive turning response before adaptation, leading to a zig-zag path of chemotaxis. In the absence of resensitization due to inhibition of protein synthesis, such asymmetry in desensitization persists after the first cycle of desensitization, leading to an overall repulsive response.

Previous studies have shown that dendrites and axons contain both messenger RNAs and the machinery for local protein translation^{45–47}. In cultured retinal ganglion cells, both netrin-1 and semaphorin 3A (Sem3A) were shown to stimulate protein synthesis in neurites severed from cell bodies, and inhibition of translation resulted in a loss of growth-cone turning induced by a gradient of netrin-1 or Sema3A⁴⁴. Our results suggest that translational inhibition affects the resensitization/adaptation of the growth cone, rather than initial growth-cone turning responses, resulting in the final repulsive turning in the netrin-1 gradient after the initial cycle of desensitization. Such repulsive turning was in fact observed for a substantial fraction of retinal ganglion axons after inhibition of translation⁴⁴. Whether long-range chemorepulsion of growth cones by gradients of repellants also requires adaptive changes in the sensitivity of the growth cone remains to be determined.

Guidance of growth cones by extracellular gradients of guidance cues involves a complex sequence of cellular events that re-adjust the sensitivity of signalling pathways to the guidance cue. Although the immediate turning of the growth cone can be achieved by rearrangement of the pre-existing cellular components, as shown by rapid asymmetric filopodia initiation within minutes after the onset of the gradient^{13,15}, successful chemotaxis of the growth cone in the gradient of guidance cues requires activation of MAPK and initiation of protein translation, a more protracted process that allows adaptive changes in the sensitivity of the growth cone. Although it remains to be determined whether adaptation is a general phenomenon of axon guidance in the developing nervous system, the existence of the adaptive behaviour in growth-cone chemotaxis described here exemplifies Ramón y Cajal's imagery of the growth cone as a chemotaking amoeba, and suggests common mechanisms underlying all forms of directed cell motility. □

Methods

Culture preparation

Cultures of *Xenopus* spinal neurons were prepared from one-day-old embryos by methods described previously^{13,15–17,48}. The experiments were carried out with 14–20 h cultures at 20–22 °C in saline containing (in mM): 140 NaCl, 2.5 KCl, 1 MgCl₂, 1 CaCl₂ and 10 HEPES, pH 7.4. For MAPK assays, explant cultures were prepared from stage 22–23 embryos. Neural tubes were treated with collagenase/dispase (0.1%, Sigma) for 10 min and triturated to isolate the spinal cord. They were then cultured in the presence of 50 µM GM6001 (Chemicon) for 14–16 h before biochemical assays (each with 10–12 explants) to ensure optimal expression of netrin receptor DCC⁴⁹.

Growth cone turning assay

Microscopic gradients of guidance factors were produced by methods described previously^{15–18}. The standard gradient used 5 or 50 µg ml^{−1} netrin-1 or BDNF, respectively, in the micropipette. Previous analyses^{15,16} have shown that, under standard pulsing conditions, the average concentration of the factor at the growth cone is about 10²-fold lower than that in the pipette, and the concentration gradient across the growth cone at a distance of 100 µm from the pipette tip is 5–10%. For the growth-cone turning assay, the tip of the pipette containing the factor was placed 100 µm away from the centre of the growth cone at an angle of 45° with respect to the initial direction of neurite extension (set by the last 10-µm segment of the neurite). The turning angle was defined by the angle between the original direction of neurite extension and a straight line connecting the positions of the centre of the growth cone at the onset and the end of the 30 min period. Only those growth cones of isolated neurons with net neurite extension greater than 5 µm over the 30-min period were included for analysis. Significance was assessed using the Kolmogorov–Smirnov test unless otherwise noted. In quantitative assays of growth-cone

adaptation after bath incubation of the guidance cue, the degradation of netrin-1 and BDNF in the bath solution over the duration of these experiments was negligible. A bath solution containing 1 ng ml^{−1} netrin-1 was fully active in causing desensitization when repeatedly used in other fresh cultures within 10 h, whereas 50% reduction in netrin-1 (0.5 ng ml^{−1}) resulted in significant reduction in desensitization (see Fig. 2a). This was further confirmed by a titration experiment: a 5 ng ml^{−1} netrin-1 bath solution was used in a desensitization experiment for 90 min. The solution was then diluted five times and examined for its effect on fresh, new cultures. Full desensitization (turning angle $-0.4 \pm 3.1^\circ$, s.e.m., $n = 11$) was found, suggesting that the degradation of netrin-1, if it occurs at all, must be less than 50%.

Calcium imaging

We carried out calcium imaging as reported previously²¹. In some experiments, the Ca²⁺ indicator Oregon green 488 BAPTA-1 conjugated to dextran (relative molecular mass of 70,000; Molecular Probes) was co-injected with Texas Red-conjugated dextran (300 µM; Molecular Probes), and Ca²⁺ imaging was done using a Leica confocal imaging system, using a $\times 63$ objective (HCX PI Apo; NA, 1.32) and separate Kr and Ar gas lasers. The Oregon green BAPTA-dextran and Texas Red-dextran were excited at 488 and 568 nm, respectively, and the respective emission signals were collected at 495–550 nm and 595–680 nm. Fluorescence images were collected sequentially in pairs every 20–30 s and analysed with NIH Image Software. We also collected transmission images to monitor laser output. Mean fluorescence intensity was measured over a fixed square area that covered the entire growth cone. Ratios of normalized Oregon green and Texas red fluorescence were used to control for fluctuations in growth-cone thickness or changes in focal plane that may have occurred during experiments.

MAPK assays

Myelin basic protein (MBP) assay was performed as previously described²⁷. Spinal cord explants were treated with netrin-1 (20 ng ml^{−1}) or BDNF (10 ng ml^{−1}) and stopped by ice cold RIPA buffer containing 50 mM Tris at pH 8, 50 mM NaF, 5 mM EDTA, 150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS, 10 µg ml^{−1} leupeptin, 50 µg ml^{−1} phenylmethylsulphonyl fluoride (PMSF), 2 µg ml^{−1} aprotinin, 1 mM ortho-vanadate and 50 µM GM6001 (Chemicon). The protein content for each sample was determined as described⁵⁰ and normalized. After clarification of the lysates with sepharose (Amersham-Pharmacia), MAPK was immuno-precipitated with 1.5 µg anti-MAPK antibody (Promega) pre-absorbed to protein A sepharose (Amersham-Pharmacia) for 2 h at 4 °C. The beads were washed twice with ice-cold RIPA and once with kinase assay buffer (KAB (in mM): 25 HEPES, pH 7.4, 25 β-glycerophosphate, 25 MgCl₂, 0.1 Na ortho-vanadate, 0.5 dithiothreitol), then incubated with 30 µl of 'hot' assay mix (KAB, 5 µg MBP, 20 µM ATP, 10 µCi [γ -³²P]ATP) for 15 min and the reaction stopped by adding 5 µl EDTA (0.5 M). Twenty-five microlitres from each reaction were subjected to SDS-polyacrylamide gel electrophoresis. The radioactive bands were exposed to X-ray film and then excised for scintillation counting. We performed immunoblot assays for MAPK according to the manufacturer's protocols (Promega).

Received 4 December 2001; accepted 8 April 2002.

Published online 1 May 2002, DOI 10.1038/nature745.

1. Tessier-Lavigne, M. & Goodman, C. S. The molecular biology of axon guidance. *Science* **274**, 1123–1133 (1996).
2. Mueller, B. K. Growth cone guidance: first steps towards a deeper understanding. *Annu. Rev. Neurosci.* **22**, 351–388 (1999).
3. Ramón y Cajal, S. *Histology of the Nervous System* (trans. Swanson, N. & Swanson, L. W.) Vol. 1, 532–537 (Oxford Univ. Press, New York, 1995).
4. Tessier-Lavigne, M., Placzek, M., Lumsden, A. G., Dodd, J. & Jessell, T. M. Chemotropic guidance of developing axons in the mammalian central nervous system. *Nature* **336**, 775–778 (1988).
5. Pini, A. Chemorepulsion of axons in the developing mammalian central nervous system. *Science* **261**, 95–98 (1993).
6. Lumsden, A. G. & Davies, A. M. Earliest sensory nerve fibres are guided to peripheral targets by attractants other than nerve growth factor. *Nature* **306**, 786–788 (1983).
7. Keynes, R. et al. Surround repulsion of spinal sensory axons in higher vertebrate embryos. *Neuron* **18**, 889–897 (1997).
8. Song, H.-J. & Poo, M.-M. The cell biology of neuronal navigation. *Nature Cell Biol.* **3**, E81–E88 (2001).
9. Baier, H. & Bonhoeffer, F. Axon guidance by gradients of a target-derived component. *Science* **255**, 472–475 (1992).
10. Rosentreter, S. M. et al. Response of retinal ganglion cell axons to striped linear gradients of repellent guidance molecules. *J. Neurobiol.* **37**, 541–562 (1998).
11. Loschinger, J., Weth, F. & Bonhoeffer, F. Reading of concentration gradients by axonal growth cones. *Phil. Trans. R. Soc. Lond. B Biol. Sci.* **355**, 971–982 (2000).
12. Falke, J. J., Bass, R. B., Butler, S. L., Chervitz, S. A. & Danielson, M. A. The two-component signalling pathway of bacterial chemotaxis: a molecular view of signal transduction by receptors, kinases, and adaptation enzymes. *Annu. Rev. Cell Dev. Biol.* **13**, 457–512 (1997).
13. Ming, G. L. et al. cAMP-dependent growth cone guidance by netrin-1. *Neuron* **19**, 1225–1235 (1997).
14. Yee, K. T., Simon, H. H., Tessier-Lavigne, M. & O'Leary, D. M. Extension of long leading processes and neuronal migration in the mammalian brain directed by the chemoattractant netrin-1. *Neuron* **24**, 607–622 (1999).
15. Lohof, A. M., Quillan, M., Dan, Y. & Poo, M.-M. Asymmetric modulation of cytosolic cAMP activity induces growth cone turning. *J. Neurosci.* **12**, 1253–1261 (1992).
16. Zheng, J. Q., Felder, M., Connor, J. A. & Poo, M.-M. Turning of nerve growth cones induced by neurotransmitters. *Nature* **368**, 140–144 (1994).
17. Song, H.-j., Ming, G.-L. & Poo, M.-M. Cyclic-AMP-induced switching in turning direction of nerve growth cones. *Nature* **388**, 275–279 (1997).
18. Ming, G., Lohof, A. M. & Zheng, J. Q. Acute morphogenic and chemotropic effects of neurotrophins on cultured embryonic *Xenopus* spinal neurons. *J. Neurosci.* **17**, 7860–7871 (1997).

19. Tucker, K. L., Meyer, M. & Barde, Y. A. Neurotrophins are required for nerve growth during development. *Nature Neurosci.* **4**, 29–37 (2001).
20. Ming, G. *et al.* Phospholipase C- γ and Phosphoinositide 3-kinase mediate cytoplasmic signalling in nerve growth cone guidance. *Neuron* **23**, 139–148 (1999).
21. Hong, K., Nishiyama, M., Henley, J., Tessier-Lavigne, M. & Poo, M. Calcium signalling in the guidance of nerve growth by netrin-1. *Nature* **403**, 93–98 (2000).
22. Kater, S. B., Mattson, M. P., Cohan, C. & Connor, J. Calcium regulation of the neuronal growth cone. *Trends Neurosci.* **11**, 315–321 (1988).
23. Gu, X. & Spitzer, N. C. Distinct aspects of neuronal differentiation encoded by frequency of spontaneous Ca^{2+} transients. *Nature* **375**, 784–787 (1995).
24. Gomez, T. M. & Spitzer, N. C. *In vivo* regulation of axon extension and pathfinding by growth-cone calcium transients. *Nature* **397**, 350–355 (1999).
25. Gomez, T. M., Robles, E., Poo, M. & Spitzer, N. C. Filopodial calcium transients promote substrate-dependent growth cone turning. *Science* **291**, 1983–1987 (2001).
26. Huang, E. J. & Reichardt, L. F. Neurotrophins: roles in neuronal development and function. *Annu. Rev. Neurosci.* **24**, 677–736 (2001).
27. Seger, R. *et al.* Overexpression of mitogen-activated protein kinase kinase (MAPKK) and its mutants in NIH 3T3 cells. Evidence that MAPKK involvement in cellular proliferation is regulated by phosphorylation of serine residues in its kinase subdomains VII and VIII. *J. Biol. Chem.* **269**, 25699–25709 (1994).
28. DeSilva, D. R. *et al.* Inhibition of mitogen-activated protein kinase kinase blocks T cell proliferation but does not induce or prevent anergy. *J. Immunol.* **160**, 4175–4181 (1998).
29. Dudley, D. T., Pang, L., Decker, S. J., Bridges, A. J. & Saltiel, A. R. A synthetic inhibitor of the mitogen-activated protein kinase cascade. *Proc. Natl Acad. Sci. USA* **92**, 7686–7689 (1995).
30. Sonenberg, N. & Gingras, A. C. The mRNA 5' cap-binding protein eIF4E and control of cell growth. *Curr. Opin. Cell Biol.* **10**, 268–275 (1998).
31. Waring, P. DNA fragmentation induced in macrophages by gliotoxin does not require protein synthesis and is preceded by raised inositol triphosphate levels. *J. Biol. Chem.* **265**, 14476–14480 (1990).
32. Kochi, S. K. & Collier, R. J. DNA fragmentation and cytolysis in U937 cells treated with diphtheria toxin or other inhibitors of protein synthesis. *Exp. Cell Res.* **208**, 296–302 (1993).
33. Mittleman, B., Zandomeni, R. & Weinmann, R. Mechanism of action of 5,6-dichloro-1-beta-D-ribofuranosylbenzimidazole. II. A resistant human cell mutant with an altered transcriptional machinery. *J. Mol. Biol.* **165**, 461–473 (1983).
34. Burhans, W. C. *et al.* Emetine allows identification of origins of mammalian DNA replication by imbalanced DNA synthesis, not through conservative nucleosome segregation. *EMBO J.* **10**, 4351–4360 (1991).
35. Martin, S. J., Lennon, S. V., Bonham, A. M. & Cotter, T. G. Induction of apoptosis (programmed cell death) in human leukemic HL-60 cells by inhibition of RNA or protein synthesis. *J. Immunol.* **145**, 1859–1867 (1990).
36. Hausdorff, W. P., Caron, M. G. & Lefkowitz, R. J. Turning off the signal: desensitization of beta-adrenergic receptor function. *FASEB J.* **4**, 2881–2889 (1990).
37. Sorkin, A. & Waters, C. M. Endocytosis of growth factor receptors. *Bioessays* **15**, 375–382 (1993).
38. Meyer-Franke, A. *et al.* Depolarization and cAMP elevation rapidly recruit TrkB to the plasma membrane of CNS neurons. *Neuron* **21**, 681–693 (1998).
39. Song, H.-J. & Poo, M.-M. Signal transduction underlying growth cone guidance by diffusible factors. *Curr. Opin. Neurobiol.* **9**, 355–363 (1999).
40. Ming, G., Henley, J., Tessier-Lavigne, M., Song, H. & Poo, M. Electrical activity modulates growth cone guidance by diffusible factors. *Neuron* **29**, 441–452 (2001).
41. Wang, Y., Liu, J. & Segall, J. E. MAP kinase function in amoeboid chemotaxis. *J. Cell Sci.* **111**, 373–383 (1998).
42. Chang, L. & Karin, M. Mammalian MAP kinase signalling cascades. *Nature* **410**, 37–40 (2001).
43. Takei, N., Kawamura, M., Hara, K., Yonezawa, K. & Nawa, H. Brain-derived neurotrophic factor enhances neuronal translation by activating multiple initiation processes. *J. Biol. Chem.* **276**, 42818–42825 (2001).
44. Campbell, D. S. & Holt, C. E. Chemotropic responses of retinal growth cones mediated by rapid local protein synthesis and degradation. *Neuron* **20**, 1013–1026 (2001).
45. Aakalu, G., Smith, W. B., Nguyen, N., Jiang, C. & Schuman, E. M. Dynamic visualization of local protein synthesis in hippocampal neurons. *Neuron* **30**, 489–502 (2001).
46. Steward, O. & Schuman, E. M. Protein synthesis at synaptic sites on dendrites. *Annu. Rev. Neurosci.* **24**, 299–325 (2001).
47. Job, C. & Eberwine, J. Localization and translation of mRNA in dendrites and axons. *Nature Rev. Neurosci.* **2**, 889–898 (2001).
48. Spitzer, N. C. & Lamborghini, J. E. The development of the action potential mechanism of amphibian neurons isolated in culture. *Proc. Natl Acad. Sci. USA* **73**, 1641–1645 (1976).
49. Pierceall, W. E. *et al.* Expression of a homologue of the deleted in colorectal cancer (DCC) gene in the nervous system of developing *Xenopus* embryos. *Dev. Biol.* **166**, 654–665 (1994).
50. Hill, H. D. & Straka, J. G. Protein determination using biconinonic acid in the presence of sulfhydryl reagents. *Anal. Biochem.* **170**, 203–208 (1988).

Acknowledgements

Chick recombinant netrin-1 was provided by M. Tessier-Lavigne. This work was supported by grants from the NIH.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to M.-m.P. (e-mail: mpoo@uclink.berkeley.edu) or G.-L.M. (e-mail: gming@biomail.ucsf.edu).

Thickness constraints on the icy shells of the galilean satellites from a comparison of crater shapes

Paul M. Schenk

Lunar and Planetary Institute, 3600 Bay Area Boulevard, Houston, Texas 77058, USA

A thin outer ice shell on Jupiter's large moon Europa would imply easy exchange between the surface and any organic or biotic material in its putative subsurface ocean^{1–4}. The thickness of the outer ice shell is poorly constrained, however, with model-dependent estimates ranging from a few kilometres^{5,6} to ten or more kilometres⁷. Here I present measurements of depths of impact craters on Europa, Ganymede and Callisto that reveal two anomalous transitions in crater shape with diameter. The first transition is probably related to temperature-dependent ductility of the crust at shallow depths (7–8 km on Europa). The second transition is attributed to the influence of subsurface oceans on all three satellites^{3,8,9}, which constrains Europa's icy shell to be at least 19 km thick. The icy lithospheres of Ganymede and Callisto are equally ice-rich, but Europa's icy shell has a thermal structure about 0.25–0.5 times the thicknesses of Ganymede's or Callisto's shells, depending on epoch. The appearances of the craters on Europa are inconsistent with thin-ice-shell models¹ and indicate that exchange of oceanic and surface material could be difficult.

The thickness of Europa's ice shell controls whether exchange between Europa's surface and its putative water ocean (including exposure of possible organic materials to sunlight¹) is possible, and in what manner. A thick shell (>10 km thick) would be susceptible to solid-state convection and overturn¹⁰, whereas a thin ice shell (only a few kilometres thick) would be vulnerable to crack-through and melt-through from below^{1,5}. Efforts to constrain the thickness of Europa's icy shell through geologic^{5–7} and geophysical^{3,4} investigations have thus far failed to produce a consensus because of controversy over the origin of geologic structures^{1,2} and weak model constraints. Impact craters (Fig. 1) sample planetary lithospheres over a range of depths and offer direct probes into Europa's icy shell. Crater morphology is influenced by planetary surface gravity¹¹ and lithospheric properties^{12,13}. Surface gravity on the three icy galilean satellites and the Moon are similar (Europa, 132 cm s^{–2}; Ganymede, 143 cm s^{–2}; Callisto, 124 cm s^{–2}; the Moon, 162 cm s^{–2}), so differences in crater morphology can be attributed to compositional or rheological differences.

Moore *et al.*^{14,15}, in their initial survey of Galileo observations of crater morphologies on Europa, found that the largest craters (Fig. 1) are morphologically distinct from and shallower than similar-sized craters on Ganymede and Callisto (Fig. 1). They concluded that these craters may have excavated into either a global subsurface ocean or a ductile zone in the lower shell, but did not conduct a systematic analysis of crater dimensions. Recent attempts to model impact melt production suggest that the ice shell may be at least 3–4 km thick¹⁶, but this does not constitute a definitive test of the 'thin crust, thick crust' debate.

Depth/diameter (d/D) data for fresh, undegraded, craters on the Moon and other planets follow well-defined trends and feature two breaks in slope^{11,17}. These breaks, or transitions, correlate with morphologic transitions from simple bowl-shaped to complex morphologies (featuring slump terraces and uplifted central peaks) and from complex craters to ringed basins (see, for example, ref. 18). For comparison, I have catalogued all craters on Ganymede and Callisto that are over 30 km across, and over 1 km across on Europa. I have also measured the depths of fresh (unmodified and

unrelaxed) craters from 100 m to 170 km in diameter on the icy galilean satellites (Fig. 2) using Voyager and Galileo images. Three transitions in crater shape have been identified on each satellite, but two of these are anomalous and occur at different diameters on Europa than on Ganymede and Callisto.

Transition I (from simple to complex morphology) is similar on all three satellites (Fig. 2) and can be associated with regular crater modification processes¹⁸. Transition II on these satellites correlates with anomalous changes in complex crater dimensions and morphology. On Ganymede and Callisto, this occurs at $D \approx 26$ km, where crater depths begin to decrease slightly ($d \approx 1.1 \pm 0.3$ km) and are shallower than predicted from extrapolation of the d/D curve for complex craters (Fig. 2a, b). This change coincides with the morphologic transition from central peak craters to central pit and dome craters¹⁹ (Fig. 1). Transition II on Europa occurs at $D \approx 8$ km, where complex craters on Europa diverge from their galilean siblings and become increasingly shallower than similar-sized complex craters on Ganymede and Callisto (Fig. 2c). Central peaks still occur but crater rim and peak morphology becomes progressively disrupted with diameter (Fig. 1).

Transition III is associated with a sharp reduction in crater depths and the development of anomalous impact morphologies. On Ganymede and Callisto, many large central dome craters have poorly developed or non-existent rims (Fig. 1). Floors of these 'anomalous dome craters' (Fig. 1) are similar to or elevated a few 100 m above surrounding plains. Rim-to-floor depths are negligible to about 500 m (Fig. 2b). Anomalous dome morphologies in older craters (based on superposed crater densities) occur at $D > 60$ km, indicating a Transition III diameter as low as 60 km in post-bright-terrain formation times. For younger central dome craters only those larger than about 150 km in diameter have anomalous morphologies, suggesting that Transition III diameters on Ganymede and Callisto have increased to about 150 km at present.

Transition III on Europa occurs at $D \approx 30$ km (Fig. 2c) and correlates with an abrupt change from modified central peak to multiring morphologies between 27 and 33 km crater diameter. The two largest impact structures, Callanish and Tyre ($D \approx 33$ and 41 km, respectively, as determined by scaling from mappable ejecta deposits²⁰) are characterized by several (>5) concentric ridges and

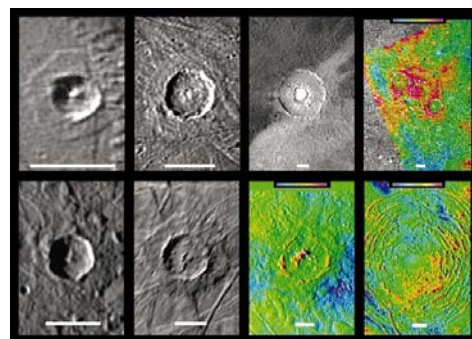


Figure 1 Impact crater landforms on the icy galilean satellites, representing the major morphologic types. In order of increasing diameter from left to right crater types are, for Ganymede and Callisto (top row): central peak craters ($D = 18$ km), central pit craters ($D = 30$ km), central dome craters (Enkidu, $D = 121$ km) and anomalous dome craters ($D = 138$ km). For Europa (bottom row): central peak craters ($D = 8$ km), modified central pit craters (Grainne, $D = 14$ km), anomalous central peak craters (Pwyll, $D = 27$ km), and multiring basins (Tyre, $D = 41$ km). Images have been scaled so that craters appear to have the same diameter. Scale bars are 30 km for Ganymede and Callisto (top row), and 10 km for Europa (bottom row). Images for Pwyll and Tyre and the unnamed Ganymede basin have been colour-coded to represent topography (red, high; blue, low; colour scale bars represent relative elevation: 0 to 500 m for Europa craters and 0 to 2,000 m for Ganymede). Note the progressively more complex central peak morphology with increasing crater diameter. The nominal rim location for Tyre lies within the innermost of the prominent topographic ridges. The Sun is from the left in all images.

graben¹⁵, but no central structures, depressions or rim scarps can be identified within the nominal crater (Fig. 1). Maximum effective crater depths, inferred from isolated ridges near the rim location and subtle topography across the floor, are between 50 and 100 m (Figs 1, 2c). These structures differ sharply from terrestrial planet multiring basins which have only 3–5 ring scarps and topography greater than 4 km (for example, refs 13, 17).

The rollover and precipitous drop in crater depths with diameter for the galilean satellites (Fig. 2) is unprecedented among the planets. That this is not due to long-term post-impact flattening

by creep (that is, viscous relaxation) is suggested by the youth of the craters studied (some have well-preserved bright rays) and by their misshapen or absent central structures and rims. Relaxation and volcanism may flatten or bury topography but are unlikely structurally to disrupt rims and central peaks (see ref. 21). Rugged relief on crater floors (Fig. 1) is also inconsistent with the idea that craters formed within a giant pool of impact melt¹⁶. Thus, differences in internal structure between the icy satellites and the Moon, and between the icy galilean satellites, are the most plausible explanation for the unusual shapes of larger craters on these satellites. Warm ice probably at shallow depths within these satellites is notoriously weaker than common silicates and the very cold brittle ice near their surfaces (see ref. 22). I propose that Transitions II and III reveal at least two such temperature-induced rheologic transitions or phase changes with depth on all three icy galilean satellites, but that these transitions are much shallower on Europa. Inherently weaker ice at greater depths would be progressively less able to support the topography associated with larger transient craters several kilometres deep²³, leading to enhanced structural collapse.

Impact into rheologically layered targets is imprecisely understood. Studies of terrestrial and lunar craters^{18,24} and recent modelling²³ of crater floor rebound suggest that modification from simple to complex morphology (by floor rebound and rim collapse) involves a quasi-hemispherical zone around the transient (pre-modification) crater¹⁸ and that the modification process and final crater shape begins to be unduly influenced when an anomalously weak layer is roughly 1 to 1.3 times as deep as the transient crater width. The McKinnon and Schenk²⁵ scaling relationship for icy satellite craters is used to estimate transient crater dimensions (to within approximately 10%) from the observed crater diameters. On Ganymede and Callisto, Transition II craters ($D \approx 26$ km) scale to transient craters about 16.5 km wide for a depth to the first anomalous layer of ~ 16 –22 km. On Europa, Transition II craters ($D \approx 8$ km) scale to transient craters about 6 km wide, for a layer depth of 7–8 km. On Ganymede and Callisto, Transition III craters ($D \approx 150$ km) scale to transient craters about 80 km wide, for a depth to the second anomalous layer of 80–105 km (the older Transition III diameter of ~ 60 km translates to transient diameters of around 35 km, and a layer depth of 35–45 km). On Europa, Transition III craters ($D \approx 30$ km) scale to transient craters about 19 km wide, and a depth to the second transition of 19–25 km.

Transition II is interpreted to represent a temperature-dependent rheologic change with depth. The preservation of prominent (albeit modified) central peaks in complex craters not yet at Transition II is consistent with solid-state material at the base of the transient crater^{14,16}. Rounded central domes, modified central peaks and degraded rim topography in craters larger than those at Transition II (Fig. 1) are all consistent with coherent but very warm ductile ice in the lower ice shell (see ref. 19).

Galileo tracking data indicate that the interior of Ganymede is probably strongly differentiated with an ice-rich lithosphere and mantle, whereas Callisto is not fully differentiated and may or may not have an ice-rich outer lithosphere^{26,27}. The indistinguishable d/D statistics for craters on Ganymede and Callisto (Fig. 2) indicate that internal composition and thermal structure of these two satellites are similar, and that Callisto also has an ice-rich outer shell at least several tens of kilometres thick.

Galileo magnetometer observations are consistent with internal oceans currently within both Ganymede and Callisto at depths of 150–200 km but possibly shallower^{8,9}. If Transition III is a detection of oceans on Ganymede and Callisto, they must be at least 80 km deep at present. Changes in morphology and transitions on Ganymede and Callisto as described here may record changes in lithospheric thermal properties²⁸ and possibly ocean depths over time.

Transition III to multiring morphologies on Europa may indicate the beginning of the influence of a global ocean¹⁵. The unusual multiring morphologies on icy satellites are probably related to

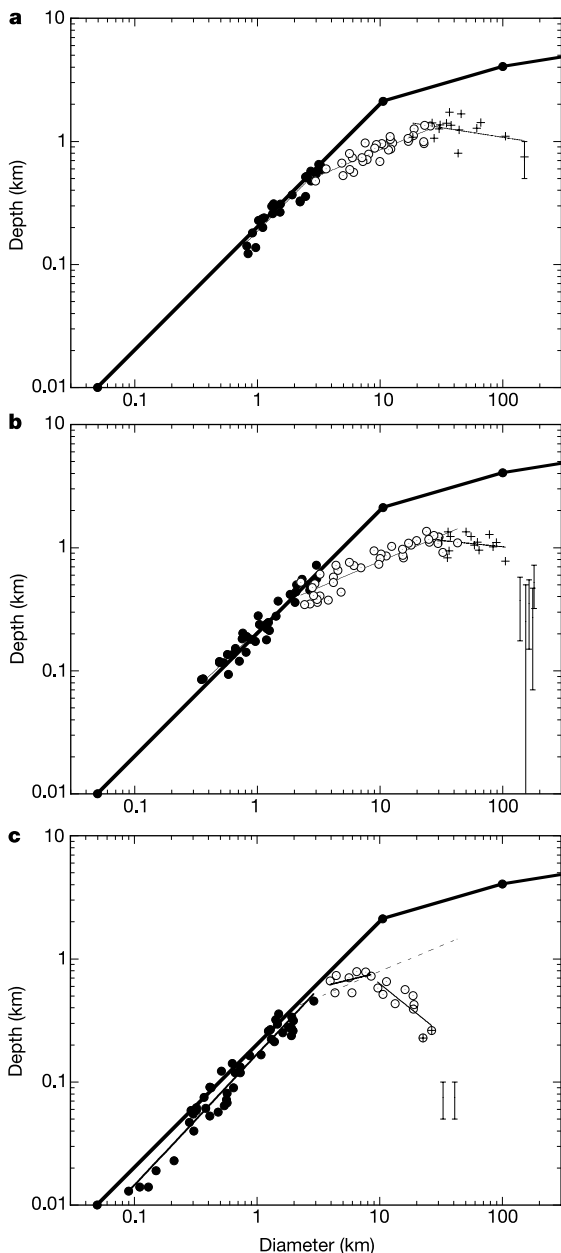


Figure 2 Depth/diameter measurements for fresh impact craters on the icy galilean satellites. Data are plotted separately for Callisto (a), Ganymede (b) and Europa (c). The heavy dark line is for lunar craters^{11,17}. The thin lines are least-squares fits through data. Craters plotted are simple craters (solid dots), complex (central peak) craters (open circles), central pit and central dome craters (crosses), and anomalous dome and multiring basins (error bars). For Europa, the modified central peak craters Mannann'an and Pwyll are plotted as circles with crosses. The dashed line in c is the least-squares fit through complex craters on Ganymede for comparison with complex craters on Europa. Note the sparser data set for Callisto. The error bar for the largest central dome crater on Callisto is typical for all plotted data.

impact into a thin brittle lithosphere overlying a fluid or a ductile layer that behaves as a fluid during crater collapse¹⁰. The apparent abruptness of Transition III on Europa (Figs 1, 2) is consistent with a sharp phase change at depth, such as an ice–liquid interface, but may also represent a rapid transition to the hot base of the ice shell where the ice is near melting and very ductile. It seems that the base of the icy shell is not much deeper, but the inferred Transition III depth at least indicates a minimum ice shell thickness of around 19–25 km. Europa's known impact craters are unlikely to excavate ocean material directly to the surface, however (unless the material is trapped in isolated shallow reservoirs). The largest crater, Tyre, excavates down to only about 3 km (~30–40% of the transient crater depth¹⁸), or around 15% of the minimum shell thickness.

Systematic mapping of impact crater shape and morphology provides a robust (and free) means of sampling icy satellite interiors. Ice rheology is dependent on temperature²² and these transition diameters provide new constraints on the comparative thermal structure of these satellites. Transitions II and III are 2–4 times shallower on Europa than on Ganymede or Callisto, indicating that the rheologic structure of Europa's outer shell is similarly thinner than on Ganymede and Callisto. As heat flow scales roughly linearly with the thickness of the stagnant non-convecting lid¹⁰ (possibly represented by Transition II), heat flow should to first order also be 2–4 times higher on Europa, depending on epoch. Transition III constrains Europa's ice shell to be at least 19–25 km thick, consistent with that required for convection to proceed within the ice shell¹⁰ and favouring diapirism of the lower shell for the origin of ovoid features and chaos terrains (refs 7, 29 and P.S. and R. Pappalardo, manuscript in preparation). The minimum shell thickness strongly supports thick ice shell models in general, and associated interpretations of geologic features^{2,7,29}; it may also indicate the depth to Europa's ocean, or at least the beginning of the hot basal layer of the floating ice shell. □

Methods

Crater depths have been measured from Galileo and Voyager data using three techniques: stereo digital elevation models (DEMs), photoclinometry DEMs, and shadow length measurements. Stereo DEMs resolve only the large craters Pwyll, Mannan'an, and Cilix (diameters 19–27 km) on Europa, and several anomalous dome craters on Ganymede. Photoclinometry in two horizontal dimensions can be used with single low-sun images to map topography from relative brightness. My photoclinometry technique includes the use of low-phase-angle images to model local albedo, thus reducing or eliminating one of the major systematic sources of error in photoclinometry. This technique is used here primarily to confirm the stereo and shadow measurements, but is the primary source of topographic data for the multiring basins Callanish and Tyre. Photoclinometry was also used to map topography across additional anomalous dome craters on Ganymede, supporting the conclusion that they have raised floors based on limited stereo data described above. I also use low-resolution controlled stereo DEMs to control the long-wavelength component of high-resolution photoclinometry DEMs. This was especially useful for Pwyll (Fig. 1). Height measurements based on triangulation of shadow lengths are used for all craters on Europa, Ganymede and Callisto smaller than ~15 km across. Depth measurement errors include systematic errors due to technique and those due to variations in topography along the rim crest. Systematic errors rarely exceed 10%. Variations in rim height approach about 100 m for large craters such as Pwyll. These data supercede the depth/diameter statistics for Ganymede and Callisto based on lower-resolution Voyager data^{12,19}; no Voyager-based measurements were possible for craters on Europa.

Received 21 January; accepted 8 April 2002.

- Greenberg, R. Tides and the biosphere of Europa. *Am. Sci.* **90**, 48–55 (2002).
- Pappalardo, R. *et al.* Does Europa have a subsurface ocean? Evaluation of the geologic evidence. *J. Geophys. Res.* **104**, 24015–24056 (1999).
- Kivelson, M. *et al.* Galileo magnetometer measurements: A stronger case for a subsurface ocean on Europa. *Science* **289**, 1340–1343 (2000).
- Zimmer, C., Khurana, K. & Kivelson, M. Subsurface oceans on Europa and Callisto: Constraints from Galileo magnetometer observations. *Icarus* **147**, 329–347 (2000).
- Greenberg, R., Geissler, P., Tufts, B. R. & Hoppa, G. Habitability of Europa's crust: The role of tidal-tectonic processes. *J. Geophys. Res.* **105**, 17551–17562 (2000).
- Hoppa, G., Tufts, B. R., Greenberg, R. & Geissler, P. Formation of cycloidal features on Europa. *Science* **285**, 1899–1903 (1999).
- Rathbun, J., Musser, G. & Squyres, S. Ice diapirs on Europa: Implications for liquid water. *Geophys. Res. Lett.* **25**, 4157–4160 (1998).
- Kivelson, M., Khurana, K. & Volwerk, M. *Icarus* (in the press).
- Kivelson, M. *et al.* Europa and Callisto: Induced or intrinsic fields in a periodically varying plasma environment. *J. Geophys. Res.* **104**, 4609–4625 (1999).

- McKinnon, W. Convective instability in Europa's floating ice shell. *Geophys. Res. Lett.* **26**, 951–954 (1999).
- Pike, R. in *Mercury* (eds Vilas, F., Chapman, C. & Matthews, M.) 165–273 (Univ. Arizona Press, Tucson, 1988).
- Schenk, P. Ganymede and Callisto: Complex crater formation and planetary crusts. *J. Geophys. Res.* **96**, 15635–15664 (1991).
- McKinnon, W. & Melosh, H. J. Evolution of planetary lithospheres: Evidence from multiring basins on Ganymede and Callisto. *Icarus* **44**, 454–471 (1998).
- Moore, J. M. *et al.* Large impact features on Europa: Results from the Galileo nominal mission. *Icarus* **135**, 127–145 (1998).
- Moore, J. M. *et al.* Large impact features on Europa: Results from the Galileo GEM mission. *Icarus* **151**, 93–111 (2001).
- Turtle, E. & Pierazzo, E. Thickness of a European ice shell from impact crater simulations. *Science* **294**, 1326–1328 (2001).
- Williams, K. & Zuber, M. Measurement and analysis of lunar basin depths from Clementine data. *Icarus* **131**, 107–122 (1998).
- Melosh, H. J. *Impact Cratering* (Oxford Press, Oxford, 1989).
- Schenk, P. Central pit and dome crater: Exposing the interiors of Ganymede and Callisto. *J. Geophys. Res.* **98**, 7475–7498 (1993).
- Schenk, P. & Ridolfi, F. Morphology and scaling of ejecta deposits on icy satellites. *Geophys. Res. Lett.* (in the press).
- Dombard, A. & McKinnon, W. Long-term retention of impact crater topography on Ganymede. *Geophys. Res. Lett.* **27**, 3663–3666 (2000).
- Durham, W., Kirby, S. & Stern, L. Creep of water ices under planetary conditions: A compilation. *J. Geophys. Res.* **102**, 16293–16301 (1997).
- Turtle, E. & Ivanov, B. Numerical simulations of crater excavation and collapse on Europa: Implications for ice thickness. *Lunar Planet. Sci. XXXIII* abstr. no. 1431 (2002).
- Melosh, H. J. & Gaffney, E. Acoustic fluidization and the scale dependence of impact crater morphology. *J. Geophys. Res.* **88**, A830–A834 (1983).
- McKinnon, W. & Schenk, P. Estimates of comet fragment masses from impact crater chains on Callisto and Ganymede. *Geophys. Res. Lett.* **22**, 1829–1832 (1995).
- Anderson, J., Lau, E., Sjogren, W., Schubert, G. & Moore, W. Gravitational constraints on the internal structure of Ganymede. *Nature* **384**, 541–543 (1996).
- Anderson, J. *et al.* Shape, mean radius, gravity field and internal structure of Callisto. *Icarus* **153**, 157–161 (2001).
- Passey, Q. & Shoemaker, E. in *Satellites of Jupiter* (ed. Morrison, D.) 379–433 (Univ. Arizona Press, Tucson, 1982).
- Collins, G., Head, J., Pappalardo, R. & Spaun, N. Evaluation of models for the formation of chaotic terrain on Europa. *J. Geophys. Res.* **105**, 1709–1716 (2000).

Acknowledgements

This work was supported by NASA Planetary Geology and Geophysics.

Competing interests statement

The author declares that he has no competing financial interests.

Correspondence and requests for materials should be addressed to the author (e-mail: schenk@pi.usra.edu).

Megagauss sensors

A. Husmann*, J. B. Betts†, G. S. Boebinger‡, A. Migliori‡, T. F. Rosenbaum* & M.-L. Saboungi‡

* The James Franck Institute and Department of Physics, The University of Chicago, Chicago, Illinois 60637, USA

† National High Magnetic Field Laboratory, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA

‡ Argonne National Laboratory, Argonne, Illinois 60439 USA, and CRMD-CNRS, Orléans, Cedex 2, France

Magnetic fields change the way that electrons move through solids. The nature of these changes reveals information about the electronic structure of a material and, in auspicious circumstances, can be harnessed for applications. The silver chalcogenides, Ag₂Se and Ag₂Te, are non-magnetic materials, but their electrical resistance can be made very sensitive to magnetic field by adding small amounts—just 1 part in 10,000—of excess silver^{1–4}. Here we show that the resistance of Ag₂Se displays a large, nearly linear increase with applied magnetic field without saturation to the highest fields available, 600,000 gauss, more than a million times the Earth's magnetic field. These characteristics of large (thousands of per cent) and near-linear response over a large

field range make the silver chalcogenides attractive as magnetic-field sensors, especially in physically tiny megagauss (10^6 G) pulsed magnets where large fields have been produced but accurate calibration has proved elusive. High-field studies at low temperatures reveal both oscillations in the magnetoresistance and a universal scaling form that point to a quantum origin^{5,6} for this material's unprecedented behaviour.

The drive to obtain the highest possible magnetic fields cuts across disciplinary boundaries, but it has particular implications for materials science. A magnetic field couples both to the orbital and the spin states of electrons in condensed matter and, if sufficiently powerful, can lift degeneracies, break symmetries, induce phase transitions, and generate new collective states such as the fractional

quantum Hall effect. At 10^6 G (100 T), the magnetic energy $g\mu_B H$ (where g is the g -factor, μ_B is the Bohr magneton, and H is the magnetic field strength) can dominate all others, even approaching the thermal energy $k_B T$ at room temperature for a bare magnetic g -factor of 2 (here k_B is Boltzmann's constant, and T is temperature). Direct-current magnets are incapable of such extreme performance, but non-destructive pulsed fields of 40–60 T are possible via capacitor discharge with pulse widths of 10^{-2} – 10^{-1} s (extended up to seconds when generator-driven)⁷, and 50–300 T has been achieved in microcoil or laser-driven geometries over 10^{-8} – 10^{-6} s timescales^{8,9}. Realizing the full scientific promise of the technology requires small, robust magnetic-field sensors with large dynamic range, rapid response times, and repeatable performance. A simple scaling with magnetic field strength and temperature is desirable both for primary calibration of the field and for the treatment of substantial spatial¹⁰ and temporal field gradients.

Our 60-T pulsed magnet provided a 16-ms-long pulse that had an exceptionally smooth inductive decay, free from high-frequency components because of the 1-MJ capacitor drive (Fig. 1a). The dual direct digital synthesizer (designed at the National High Magnetic Field Laboratory) and digital lock-in amplifier (LGK Corp.) was specially constructed to extract cleanly phase-sensitive voltages with uncommon accuracy over the entire field range during each magnet 'shot' (Fig. 1b). We plot in Fig. 2 the magnetic-field dependence up to 55 T of both the diagonal and off-diagonal components of the resistivity of $\text{Ag}_{2+\delta}\text{Se}$ from room temperature down to pumped helium temperature. The current was applied normal to the magnetic field direction. The normalized change in the longitudinal resistivity, $(\rho_{xx}(H) - \rho_{xx}(H=0))/\rho_{xx}(H=0) = \Delta\rho_{xx}/\rho_0$ (Fig. 2a), increases fivefold by $H = 55$ T at $T = 290$ K and almost 25-fold at $T = 1.5$ K. It evolves from a slightly superlinear to a slightly sub-linear field dependence below $T = 110$ K. At the lowest temperatures an oscillatory component of the magnetoresistance emerges, but at no temperature is there any indication that the response is

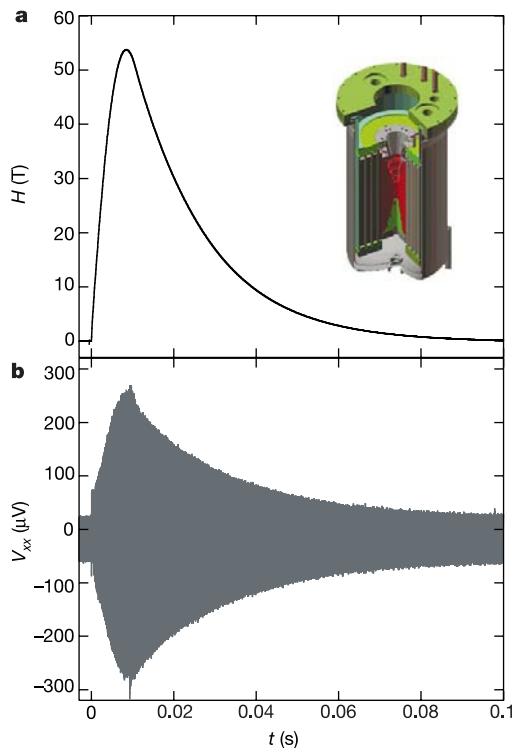


Figure 1 Pulsed magnet characteristics and low-noise data extraction. **a**, Capacitor-driven pulsed magnetic field profile as a function of time. Inset, diagram of the generator-driven 60-T long-pulse magnet, capable of providing up to 100 ms at peak field and a prime candidate for $\text{Ag}_{2+\delta}\text{Se}$ field sensors. It is about 2 m tall, 1 m in diameter, and features a 32-mm-diameter room-temperature bore. When energized it contains about 90 MJ of energy in the magnetic field. The windings of the electromagnet (shown red) are supported by cylindrical shells of high-strength, high-fracture-toughness steel. **b**, Data from a magnet 'shot'. A precision, low-distortion sine wave excitation drives a conventional 5-wire measurement of the longitudinal (V_{xx}) and Hall (V_{xy}) voltages at 156.3 kHz. A clock signal for a 16-bit analog-to-digital (A/D) converter is generated at an integer multiple of the excitation frequency. Raw data is obtained with nanosecond time jitter and megahertz raw Nyquist bandwidth. By setting the A/D clock at an integer multiple of the excitation, multiplying the A/D readings by software-generated sine and cosine functions, and averaging, a result similar to that from a commercial lock-in amplifier is produced, but with three key differences important for this experiment. First, we can adjust key measurement parameters after the magnet shot. We implement time-domain filtering of spikes, frequency-domain filtering of noise, adjust the phase, and adjust the averaging time constant after the shot and reversibly, tailoring the system response to the extreme pulsed magnet environment without having to retake data. Second, this method of synchronous digitization produces settling of the averaging 'filter' in exactly a half-cycle of excitation, thereby providing a response time that is much shorter than can be obtained with a conventional lock-in having an RC time constant final filter. This is essential for a low-noise measurement during a 16-ms experiment. Third, we can mathematically adjust the filter algorithm time centre such that there is no delay between a reported data point and the plotted magnetic field.

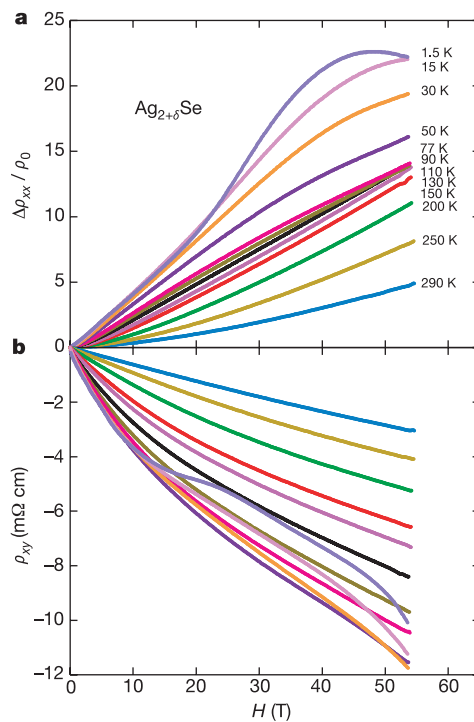


Figure 2 Magnetotransport of $\text{Ag}_{2+\delta}\text{Se}$ with $\delta \approx 10^{-4}$ in a 55-T pulsed magnetic field. **a**, The magnetoresistance normalized to its $H = 0$ value continues to climb over the full extent of the magnetic pulse at all temperatures T . **b**, Field variation of the Hall resistivity at the same temperatures (colour coded). In the low- T , low- H limit, electron density $n = 1.1 \times 10^{18} \text{ cm}^{-3}$.

saturating. Semiclassically, the orbital magnetoresistance is controlled by the product of the cyclotron frequency (ω_c) and the scattering time (τ), with a positive, quadratic magnetoresistance expected to saturate for $\omega_c\tau \approx 1$. In $\text{Ag}_{2+\delta}\text{Se}$, the quadratic behaviour required by symmetry considerations exists only below 0.01 T and the positive, approximately linear magnetoresistance continues to climb even for $\omega_c\tau > 50$, with no cut-off length scale in sight. The Hall resistivity, $\rho_{xy}(H)$ (Fig. 2b), demonstrates the expected n-type character for Ag-rich material. The crossover from intrinsic, small-gap semiconducting behaviour at high temperature (with a temperature-dependent carrier density gleaned from $d\rho_{xy}/dH|_{H \rightarrow 0}$) to a constant electron density $n = 1.1 \times 10^{18} \text{ cm}^{-3}$ at low temperature occurs at $T_p \approx 80 \text{ K}$.

After digital signal processing of the 156-kHz signal, the point-to-point noise in $\rho_{xx}(H)$ and $\rho_{xy}(H)$ is below 5 parts in 10,000, corresponding to a magnetic field uncertainty of less than 0.005 T at room temperature and $\sim 0.001 \text{ T}$ by 30 K over the entire field range. This markedly outperforms the present practice of determining H by integrating the voltage induced in a small pick-up coil inside the pulsed magnet. The small signal from such coils can be unreliable for H below a few tesla and, in a microcoil geometry, can introduce measurement uncertainties of $\sim 12\%$ at $H = 50 \text{ T}$ (ref. 8). Moreover, a magnetoresistive sensor provides an absolute calibration; there is no drift or accumulated error from integrating over time. Finally, in contrast to Hall bar sensors, these can be configured as two-lead devices that are intrinsically lower-power devices: $\rho_{xx}(H)$ measurements on a chip of silver selenide $0.4 \times 0.4 \times 1.0 \text{ mm}$ dissipate less than 10 nW of power.

New techniques are being developed to exploit advances in high field capability, from refined medical imaging and high-field NMR spectroscopy for protein analysis, to picosecond measurements in a.c. fields and synchrotron X-ray scattering in pulsed, single-turn coils. Corresponding developments in magnet sensor technology are required. The state of the art for d.c. magnets is the Hall sensor, but at the cost of large power dissipation owing to high impedances. Large RC time constants (where R is resistance and C is capacitance) preclude any high-frequency applications. Pulsed magnets depend on pick-up coils, fine wire typically wound in 10–100-turn coils, sensitive to dH/dt . Calibration errors of 5–10% can be reduced

by adding more turns n , but this increases the L/R time constant (as n^2/n). (Here L is inductance.) Furthermore, the highest-field magnets have the smallest bores; smaller-volume pick-up coils have more-limited accuracy. The silver chalcogenides are uniquely suited to both d.c. and a.c. fields. They offer the important freedom of measuring $\rho_{xx}(H, T)$ at any suitably high frequency, avoiding noisy frequencies in a given application (for example, frequencies used in the radio-frequency coil of an MRI (magnetic resonance imaging) magnet) and minimizing crosstalk. They can be shrunk in size without reducing total signal, and placed near or upon the sample as magnet bore sizes, and field homogeneity, decrease.

We now consider if it is possible to collapse the data of Fig. 1 onto a universal curve that describes the behaviour of the material at all fields and temperatures. In conventional metals, the electrical conductivity is described in terms of the Boltzmann equation, assuming semiclassical transport with a single scattering time τ . In this framework, the magnitude of the magnetic field only enters into the expression for the normalized change in resistivity $\Delta\rho_{xx}/\rho_0$ in the form of the product $H\tau$. As $\tau^{-1} \propto \rho_0$, $\Delta\rho_{xx}/\rho_0$ can be expressed solely as a function of H/ρ_0 . This functional dependence is known as Kohler's rule¹¹, and it is obeyed in a great number of metals where ρ_0 is tuned by impurity density or by temperature¹². Kohler's rule is also a powerful test of whether the semiclassical approach can describe magnetotransport in exotic materials like the organic conductors¹³ and the copper oxide superconductors¹⁴. In semiconductors, the carrier density n may vary when ρ_0 is tuned by changing T ; a modified Kohler plot then scales $\Delta\rho_{xx}/\rho_0$ with $H/n\rho_0$.

At the extreme fields of our experiment, the Hall resistivity ρ_{xy} of semiconducting $\text{Ag}_{2+\delta}\text{Se}$ is no longer strictly linear in H (Fig. 2b), and n becomes a function of H . In order to account for this added dependence, we proceed by scaling $\Delta\rho_{xx}/\rho_0$ with $|\rho_{xy}|/\rho_0$ for $1.5 < T < 290 \text{ K}$. $\Delta\rho_{xx}$ and ρ_{xy} are taken at the same H which, like n , enters only implicitly. The scans at each temperature are straight lines that lie parallel to each other on a log-log plot (not shown), indicating that one power law can serve as the general Kohler-type function. The curves do not coincide, but are shifted by a multiplicative factor that depends on T . Collapse onto a universal curve over almost three decades in $\Delta\rho_{xx}/\rho_0$ is demonstrated in Fig. 3 using the scaling form: $b(T)\Delta\rho_{xx}/\rho_0 = f(\rho_{xy}/\rho_0)$. The plot covers the field range from 55 T down to 0.5 T, below which the limitations of the pick-up coil used to calibrate our pulsed magnets introduce offsets and scatter into the data. Data for $T < 30 \text{ K}$ oscillate about the universal curve and are not shown. Magnetoresistance and Hall data from a second, more heavily doped $\text{Ag}_{2+\delta}\text{Se}$ sample similarly collapse onto a universal curve consistent with the same power law

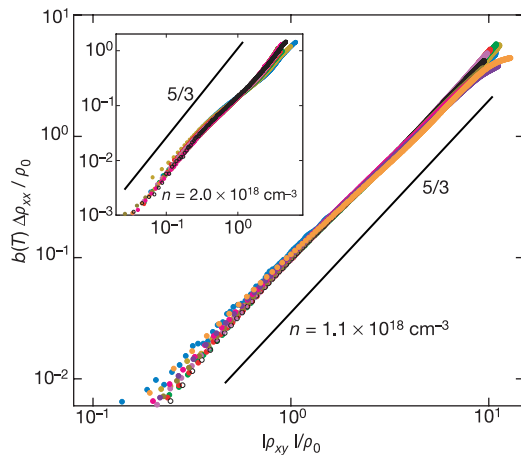


Figure 3 Scaling of the data of Fig. 2 using a modified Kohler plot where both $n(H)$ and H are implicit variables. The normalized magnetoresistivity ($\Delta\rho_{xx}/\rho_0$) scales with the normalized Hall resistivity ($|\rho_{xy}|/\rho_0$) as a simple power law. The data collapse onto a universal curve for $0.5 \text{ T} < H < 55 \text{ T}$ and $30 \text{ K} < T < 290 \text{ K}$ when shifted by a multiplicative factor $b(T)$, where we have chosen $T = 290 \text{ K}$ as the baseline. The line follows $\Delta\rho_{xx} \propto \rho_{xy}^{5/3}$. Lower-temperature measurements exhibit quantum oscillations about the scaling curve and are not shown. Inset, collapse for a more highly doped sample of $\text{Ag}_{2+\delta}\text{Se}$ with carrier density n for $0.5 \leq H \leq 55 \text{ T}$ using the identical scaling form. Temperatures are 30, 50, 77, 87, 100, 150, 200, 250 and 290 K.

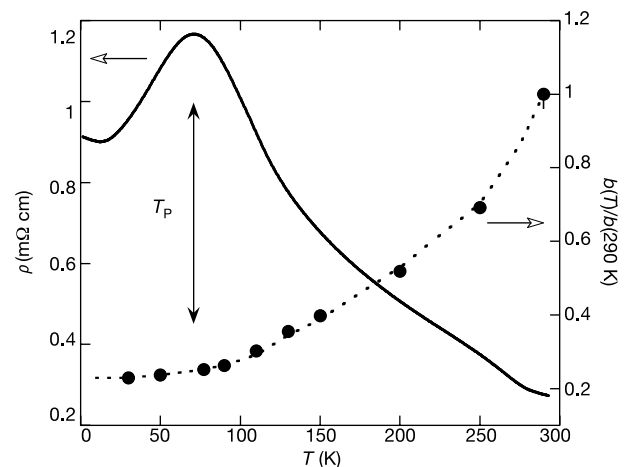


Figure 4 The scaling factor $b(T)$ and the longitudinal resistivity $\rho_{xx}(H = 0)$ as functions of temperature, suggesting different scaling character in the extrinsic and intrinsic semiconducting regimes (separated by the peak in ρ_{xx} at T_p).

(slope) and comparable absolute values for $30\text{ K} \leq T \leq 290\text{ K}$ and $0.5\text{ T} \leq H \leq 55\text{ T}$ (Fig. 3, inset). The quantum oscillations at low T are more pronounced at higher carrier density, and suggest that reducing n could extend the temperature range over which the monotonic power law scaling holds.

The $5/3$ scaling power ($\Delta\rho_{xx} \propto \rho_{xy}^{1.65 \pm 0.1}$) reflects more than the nonlinear H -dependence of the Hall resistivity (Fig. 2). Scaling works in the low-field regime where $\rho_{xy} \propto H$ and in the high-field regime where $\rho_{xy}(H)$ departs from linearity; hand-in-hand variations of $\rho_{xx}(H, T)$ and $\rho_{xy}(H, T)$ naturally conspire to produce power-law scaling over the full range of field and temperature. Collapse onto one curve (Fig. 3), however, requires the sample-specific factor $b(T)$, whose general form is illustrated in Fig. 4. We index $b(T)$ to $\rho_{xx}(H=0, T)$, noting the distinction between the constant carrier density regime below the resistance peak and the activated regime for $T > 80\text{ K}$ (ref. 1). The fact that $b(T)$ becomes markedly less T -dependent below $T \approx 80\text{ K}$ argues for a decomposition of $b(T)$ above and below T_p . At low T , ρ_{xx} increases with increasing T owing to phonon scattering, $b(T)$ is a nearly temperature-independent constant, and Kohler's rule serves as a reasonable construct for seeking universality. Complications ensue in the intrinsic semiconducting regime above T_p . We have nominally accounted for the variation of n with T and H by plotting $\Delta\rho_{xx}/\rho_0$ versus ρ_{xy}/ρ_0 rather than $H/n\rho_0$. Nonetheless, $b(T)$ rises steadily as $\rho_{xx}(T)$ falls, pointing to elements beyond semiclassical transport.

A quantum transport mechanism is certainly suggested by the pronounced oscillations in $\rho_{xx}(H)$ and $\rho_{xy}(H)$ that emerge at low T . They begin for $\rho_{xx} \approx 0.8\text{ m}\Omega\text{ cm}$, which corresponds to a sheet resistance (per unit cell layer) of $14\text{ k}\Omega$, close to the fundamental unit of resistance given by h/e^2 (where h is Planck's constant and e is the electronic charge). The oscillatory minima do not scale precisely as $1/H$, indicating a field-induced change in the electronic density of states beyond simple Landau level formation. Quantum fluctuations combined with disorder, caused for example by excess silver embedded in less-conducting material, can introduce length scales that are not set by the cyclotron radius, and can lead to a linear field dependence of the resistivity over decades in H (refs 5 and 6). (Classical approaches to inhomogeneous media¹⁵ also can account for the unusual linear magnetoresistance in non-stoichiometric Ag_2Se and Ag_2Te , but are, at present, incompatible with experiments at small $\omega_c\tau$; refs 1, 4.)

The technological promise of the silver chalcogenides lies precisely in the smooth and continued development of the magnetotransport characteristics over huge changes in applied field. Perhaps most telling, however, is the robust scaling relation that unfolds while relative energy scales change by orders of magnitude. The universal behaviour of Fig. 3 holds when the magnetic energy $g\mu_B H$ is small compared to all others, and when it rivals the gap energy¹⁶, the thermal energy, and the energy scale set by the temperature dependence of the chemical potential. The $5/3$ scaling exponent cannot be derived from classical transport equations, and may reflect a nonlinear coupling between different energy scales, common to hydrodynamic systems¹⁷. Whatever the underlying physical mechanism, the observed universal behaviour over decades in field and temperature permits predictive power for $\text{Ag}_{2+\delta}\text{Se}$ sensor response, with every expectation of magnetic field sensitivity well beyond 10^6 G . □

Methods

High-purity Ag_2Se was ground and loaded into outgassed fused silica ampoules inside a helium glove-box, and appropriate amounts of silver were added to reach the desired compositions. The samples were sealed under vacuum, heated in a rocking furnace above its melting point for 24 h, and left to cool in a horizontal position. Regularly shaped sensors of millimetre dimensions were cut on a diamond saw. Electrical leads could be attached easily with either silver epoxy or ultrasonically soldered InBi contacts. Rapid thermal cycling between liquid helium and room temperature can lead to initial changes of the resistance up to 5%, but sample characteristics stabilize after 5 to 7 quenches with no further changes. The variations appear to be dominated by shifts in the relative positions

and quality of the electrical contacts. Lithographically defining micrometre-sized contact pads should improve reproducibility, and would permit the fabrication of submillimetre-sized sensors. Encapsulation in any thermally matched material (such as epoxy) for protection should not hinder performance.

Received 17 December 2001; accepted 2 April 2002.

1. Xu, R. *et al.* Large magnetoresistance in non-magnetic silver chalcogenides. *Nature* **390**, 57–60 (1997).
2. Chuprakov, I. S. & Dahmen, K. H. Large positive magnetoresistance in thin films of silver telluride. *Appl. Phys. Lett.* **72**, 2165–2167 (1998).
3. Ogorelec, Z., Hamzic, A. & Basletic, M. On the optimization of the large magnetoresistance of Ag_2Se . *Europhys. Lett.* **46**, 56–61 (1999).
4. Schnyders, H. S., Saboungi, M.-L. & Rosenbaum, T. F. Magnetoresistance in n- and p-type Ag_2Te : Mechanisms and applications. *Appl. Phys. Lett.* **76**, 1710–1712 (2000).
5. Abrikosov, A. Quantum magnetoresistance. *Phys. Rev. B* **58**, 2788–2794 (1998).
6. Abrikosov, A. Quantum linear magnetoresistance. *Europhys. Lett.* **49**, 789–793 (2000).
7. Boebinger, G. S., Lacerda, A. H., Schneider-Muntau, H. J. & Sullivan, N. The National High Magnetic Field Laboratory's pulsed magnetic field facility in Los Alamos. *Physica B* **294–295**, 512–518 (2001).
8. Mackay, K., Bonfim, M., Givord, D. & Fontaine, A. 50 T pulsed magnetic fields in microcoils. *J. Appl. Phys.* **87**, 1996–2002 (2000).
9. von Ortenberg, M. *et al.* The Humboldt high magnetic field center at Berlin. *Physica B* **294–295**, 568–573 (2001).
10. Drndic, M., Johnson, K. S., Thywissen, J. H., Prentiss, M. & Westervelt, R. M. Micro-electromagnets for atom manipulation. *Appl. Phys. Lett.* **72**, 2906–2908 (1998).
11. Kohler, M. Zur magnetischen Widerstandsänderung reiner Metalle. *Ann. Phys.* **32**, 211–218 (1938).
12. Pippard, A. P. *Magnetoresistance in Metals* (Cambridge Univ. Press, Cambridge, 1989).
13. McKenzie, R. H., Qualls, J. S., Han, S. Y. & Brooks, J. S. Violation of Kohler's rule by the magnetoresistance of a quasi-two-dimensional organic metal. *Phys. Rev. B* **57**, 11854–11857 (1998).
14. Harris, J. M. *et al.* Violation of Kohler's rule in the normal state magnetoresistance of $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ and $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$. *Phys. Rev. Lett.* **75**, 1391–1394 (1995).
15. Herring, C. Effect of random inhomogeneities on electrical and galvanomagnetic measurements. *J. Appl. Phys.* **31**, 1939–1953 (1960).
16. Aliev, S. A. & Aliev, F. F. Band parameters and energy structure of $\beta\text{-Ag}_2\text{Se}$. *Izv. Akad. Nauk SSSR Neorg. Mater.* **21**, 1869–1872 (1985).
17. Tritton, D. J. *Physical Fluid Dynamics* (Van Nostrand Reinhold, New York, 1977).

Acknowledgements

We thank H. Hwang, L. P. Kadanoff and H. S. Schnyders for discussions, and the late R. Xu for technical assistance. The work at the University of Chicago and at Argonne National Laboratory was supported by DOE Basic Energy Sciences.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to T.F.R. (e-mail: t-rosenbaum@uchicago.edu).

Rapidly recovering hydrogel scaffolds from self-assembling diblock copolypeptide amphiphiles

Andrew P. Nowak^{†‡}, Victor Breedveld^{†‡}, Lisa Pakstis[§], Bulent Ozbas[§], David J. Pine^{†‡}, Darrin Pochan[§] & Timothy J. Deming^{*†}

^{*} Departments of Materials and Chemistry; [†] Materials Research Laboratory; [‡] Department of Chemical Engineering, University of California, Santa Barbara, California 93106, USA
[§] Department of Materials Science and Engineering and Delaware Biotechnology Institute, University of Delaware, Newark, Delaware 19716, USA

Protein-based hydrogels are used for many applications, ranging from food and cosmetic thickeners to support matrices for drug delivery and tissue replacement^{1–3}. These materials are usually prepared using proteins extracted from natural sources, which can give rise to inconsistent properties unsuitable for medical applications⁴. Recent developments have utilized recombinant DNA methods to prepare artificial protein hydrogels with specific association mechanisms and responsiveness to various stimuli^{5,6}. Here we synthesize diblock copolypeptide amphiphiles containing charged and hydrophobic segments. Dilute solutions of these

copolypeptides would be expected to form micelles; instead, they form hydrogels that retain their mechanical strength up to temperatures of about 90 °C and recover rapidly after stress. The use of synthetic materials permits adjustment of copolymer chain length and composition, which we varied to study their effect on hydrogel formation and properties. We find that gelation depends not only on the amphiphilic nature of the polypeptides, but also on chain conformations— α -helix, β -strand or random coil. Indeed, shape-specific supramolecular assembly is integral to the gelation process, and provides a new class of peptide-based hydrogels with potential for applications in biotechnology.

Diblock copolypeptide amphiphiles were synthesized (Table 1) using transition metal-mediated α -amino-acid N-carboxyanhydride polymerizations, which allow control over polypeptide chain length and composition (see Methods)^{7,8}. The poly(L-lysine-HBr) and poly(L-glutamate sodium salt) domains are highly charged polyelectrolytes at neutral pH and dissolve readily in water⁹. The hydrophobic domains, when sufficiently large, can adopt regular conformations that aggregate and are insoluble in water, namely, rod-like α -helices for poly(L-leucine) and crystalline β -sheets for poly(L-valine)⁹. Many of these diblock amphiphiles were found to form rigid hydrogels when attempts were made to dissolve small quantities (such as 0.25–2.0 wt%) in water. Copolypeptides of identical compositions, but of statistically random sequences, were found never to form hydrogels at all. Most diblock copolymer amphiphiles with narrow chain length distributions (CLD) form micelles and vesicles in water, resulting in free-flowing suspensions^{10–14}.

Micelles and vesicles do not form interconnected networks, and elastic gel properties only arise when the concentration of particles is sufficiently high to pack them into ordered arrays, typically at least 5–10 wt%, which is an order of magnitude larger than in our polypeptide samples. Block copolymer hydrogel formation at low concentrations has only been observed with triblock architectures, where associating domains on the ends of the copolymers act as interchain crosslinks that result in network formation¹⁵. Similarly, biopolymer gels such as gelatin and carrageenan have several association sites along their long polymer backbones¹⁶. Our copolypeptides displayed quite unusual behaviour in that they were assembling to form network structures from short, narrow CLD diblock amphiphilic chains at low concentrations.

To quantitatively investigate the nature of gel formation in these samples, amphiphiles with different amino acids, hydrophilic to hydrophobic ratios and a range of overall lengths were analysed using rheology¹⁷. The results are summarized in Fig. 1a, where the ratio of storage modulus to loss modulus (G'/G'') was measured for

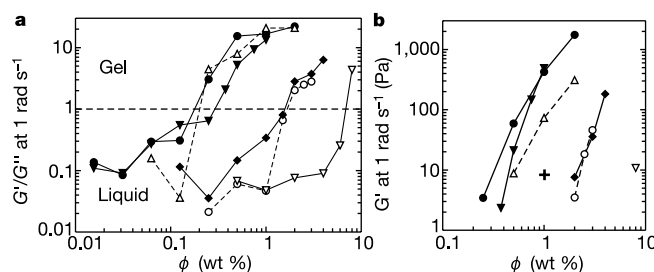


Figure 1 Gelation behaviour of aqueous diblock copolypeptide solutions. **a**, G'/G'' versus concentration ϕ for all samples, at $\omega = 1 \text{ rad s}^{-1}$ ($\phi = [(\text{grams of polymer/grams of sample}) \times 100\%]$). Filled circles, $K_{160}L_{40}$; open circles, $K_{180}L_{20}$; filled down triangles, $K_{160}V_{40}$; open down triangles, $K_{180}V_{20}$; filled diamonds, $K_{160}(\text{rac-L})_{40}$; and open up triangles, $E_{160}L_{40}$; data points are linked to guide the eye. **b**, Gel strength, represented by G' at $\omega = 1 \text{ rad s}^{-1}$, versus concentration for all samples; legend same as in **a** but with crosses indicating $K_{160}L_{40}$ + 25 mM NaCl.

different polypeptides as a function of concentration. The crossover point where $G' = G''$ provides an accurate measure for the gelation threshold, marking the transition from predominantly viscous to elastic properties. In addition to the ratio G'/G'' , the absolute strengths (G') of the gelled samples are presented in Fig. 1b. For comparison, a 2.0-wt% aqueous gelatin gel was found to possess a strength of $G' = 20 \text{ Pa}$. Because G' for all gels was observed to plateau with increased frequency (see data for $K_{160}L_{40}$ in Fig. 2a), the value at this particular frequency was a good measure of gel strength.

Contrary to most protein gels, which generally dissolve at temperatures greater than 60 °C (ref. 16), these polypeptide gels showed high thermal stability by not visibly thinning up to temperatures as high as 90 °C. Replacement of cationic poly(lysine) segments in $K_{160}L_{40}$ with anionic poly(glutamate) segments ($E_{160}L_{40}$) did not affect hydrogel formation (Fig. 1a). Addition of salt (for example 25 mM NaCl) was found to weaken the gels (Fig. 1b), presumably through screening of polyelectrolyte charges; thus showing that the highly charged segments contribute to gel formation. Accordingly, a similar copolypeptide ($K_{160}^PL_{40}$; where K^P is $\epsilon\text{-CH}_3\text{O}(\text{CH}_2\text{CH}_2\text{O})_2\text{CH}_2\text{C}(\text{O})\text{-L-lysine}$) containing non-charged hydrophilic residues in place of lysine¹⁸, was found not to form aqueous gels at concentrations up to 10 wt%.

The rheology data also revealed trends relating gelation to molecular parameters. The amphiphile $K_{180}L_{20}$ formed gels at concentrations above 2.0 wt%, yet $K_{180}V_{20}$ only formed gels at concentrations above 8.0 wt%. Because L-leucine and L-valine are similarly hydrophobic, the differences in gelation ability might be

Table 1 Abbreviations for diblock copolypeptide amphiphiles

Sample	R^1	R^2	Proposed conformation
K_xL_y	$-(\text{CH}_2)_4\text{NH}_3^+ \text{ Br}^-$	$-\text{CH}_2\text{CH}(\text{CH}_3)_2$	
K_xV_y	$-(\text{CH}_2)_4\text{NH}_3^+ \text{ Br}^-$	$-\text{CH}(\text{CH}_3)_2$	
$K_x(\text{rac-L})_y$	$-(\text{CH}_2)_4\text{NH}_3^+ \text{ Br}^-$	$-\text{CH}_2\text{CH}(\text{CH}_3)_2$	
E_xL_y	$-\text{CH}_2\text{CH}_2\text{COO}^- \text{ Na}^+$	$-\text{CH}_2\text{CH}(\text{CH}_3)_2$	

R^1 and R^2 refer to amino-acid side-chains (Fig. 3a). K, lysine; L, leucine; V, valine; E, glutamic acid. x, y are number of residues in polymer segments.

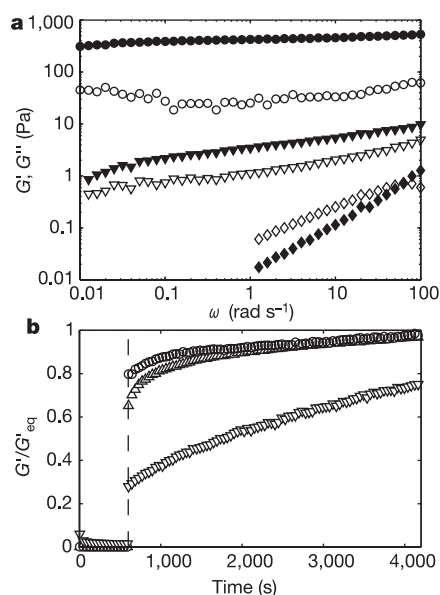


Figure 2 Rheology of diblock copolypeptide solutions. **a**, G' (ω) (filled symbols) and G'' (ω) (open symbols) of $K_{160}L_{40}$ at different concentrations: Diamonds, 0.125 wt%; down triangles, 0.25 wt%; and circles, 1 wt%. The crossover of G' and G'' in the 0.125-wt% sample is an artefact attributable to limitations of the measuring geometry (gap loading limit) and is thus not a relaxation time. **b**, Recovery of gel strength G' for: open up triangle,

1.0-wt% $K_{160}L_{40}$; open circle, 0.75-wt% $K_{160}V_{40}$; open down triangle, 2.0-wt% gelatin. Large-amplitude oscillatory breakdown ($1,000\%$ at 6 rad s^{-1} for 600 s) was followed by linear recovery measurements (0.3–1.0% at 6 rad s^{-1}). G' is normalized to the equilibrium gel strength G'_{eq} at 6 rad s^{-1} (see, for example, **a**) to facilitate sample comparison.

due to the conformational differences of the hydrophobic segments. Upon increasing the hydrophobe content to 20 mol%, both $K_{160}L_{40}$ and $K_{160}V_{40}$ were able to form hydrogels at concentrations above 0.4 wt%, with $K_{160}L_{40}$ gelling at a slightly lower concentration than $K_{160}V_{40}$. In these samples, both hydrophobic domains were found

to adopt their expected ordered conformations (Fig. 3c, d), suggesting that both the α -helical and β -strand structures promote gelation. The marked loss of β -structure in going from $K_{160}V_{40}$ to $K_{180}V_{20}$, compared to the negligible change in α -helix content in going from $K_{160}L_{40}$ to $K_{180}L_{20}$ (Fig. 3c, d), correlated well with their

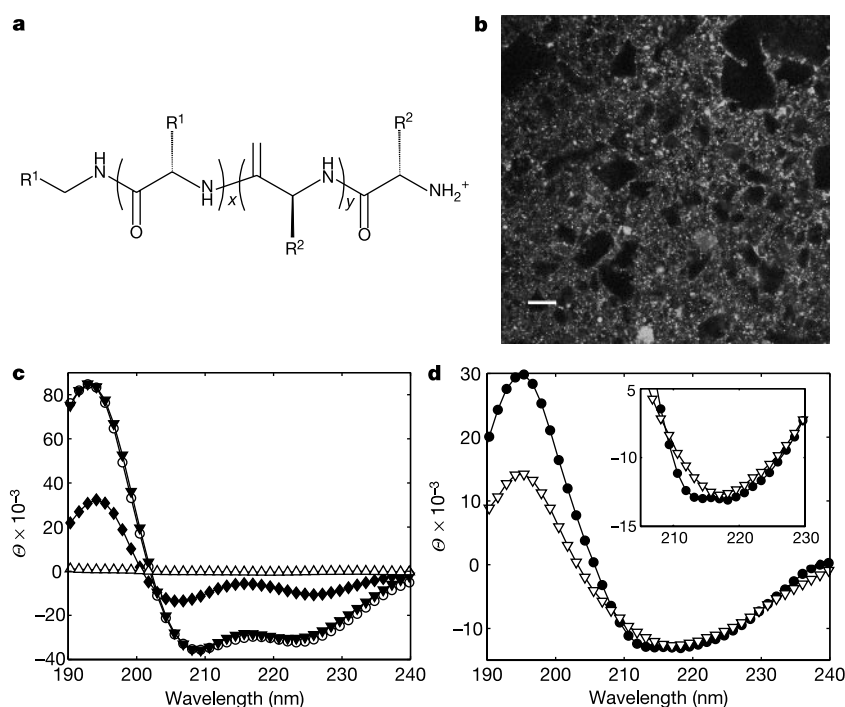


Figure 3 Molecular conformation of diblock copolypeptides. **a**, Generalized molecular structure for all samples (see Table 1 for side groups). **b**, Confocal image of 1-wt% $K_{160}L_{40}$ gel; scale bar, 20 μm . **c**, Circular dichroism spectra (molar ellipticity, Θ , in degrees $\text{cm}^2 \text{dmol}^{-1}$) of leucine-containing samples: Filled down triangle, (rac-K) $_{160}L_{40}$;

open circle, (rac-K) $_{180}L_{20}$; diamond, (rac-K) $_{90}L_{10}$; and open up triangle, (rac-K) $_{160}(\text{rac-L})_{40}$. **d**, Circular dichroism spectra of valine-containing samples: Filled circle, (rac-K) $_{180}V_{20}$; and open up triangle, (rac-K) $_{160}V_{40}$. Inset shows expanded detail of spectra.

respective decreases in gel-forming ability. Similarly, a copolypeptide of lower molecular weight, $K_{90}L_{10}$, where the helical conformation of the leucine domain was now greatly disordered because of its short length (Fig. 3c), also showed no tendency to form a gel up to concentrations as high as 8.0 wt%. These results suggest that stable chain conformations are essential in the formation and strength of these hydrogels.

To examine chain conformation explicitly we used rheological analysis of a $K_{160}(\text{rac-L})_{40}$ copolypeptide where the poly(leucine) segment contained a statistical sequence of equal proportions of L- and D-leucine residues. This hydrophobic segment, although compositionally identical to poly(L-leucine), cannot adopt the α -helical conformation because of unfavourable side-chain interactions (Fig. 3c)⁹. The substitution of this conformationally random segment resulted in a significant increase of gelation threshold (0.25 to 2.0 wt%; Fig. 1a). Therefore, it appeared that gelation is tied to the conformations of the hydrophobic domains, where α -helical segments are slightly better gelators than β -strands, which are better than random coils.

This correlation suggests that the self-assembly process is conformation-specific at the molecular level, similar to protein assembly. The specificity of these interactions was further illustrated by mixing equal amounts of $K_{160}L_{40}$ and $K_{160}(\text{rac-L})_{40}$ such that the final concentrations of each amphiphile were 0.25 wt%. If these amphiphiles could cooperate, the total amount of copolymer (0.50 wt%) should have resulted in a rigid hydrogel. However, the G' measured for this sample was lower than that of a 0.25-wt% solution of $K_{160}L_{40}$ alone. This result indicated that $K_{160}(\text{rac-L})_{40}$ was not only unable to cooperate in hydrogel formation, but also disrupted packing of the helical chains in $K_{160}L_{40}$. Similarly, chain length was found to be important in the specificity of assembly because an analogous mixture of $K_{160}L_{40}$ and $K_{180}L_{20}$ —in effect, a simulated increase in hydrophobic CLD—was found to lower G' relative to the individual components.

Although the rheological studies were sensitive to changes in macroscopic structure, they did not reveal details about the underlying gel microstructure. To study their morphology, the gels were visualized using cryogenic transmission electron microscopy (CTEM) and laser scanning confocal microscopy (LSCM), where a small amount of fluorescent hydrophobic dye was used to label the self-assembled hydrophobic matrix. In Fig. 3b and in the Supplementary Information, images are shown of a typical gel ($K_{160}L_{40}$) at concentrations matching those of rheology measurements in Fig. 2a. As concentration was increased from the free-flowing 0.125-wt% solution to the 1.00-wt% hydrogel, the images showed a progression from loosely connected fragments of dyed polypeptide-containing domains to a dense microscopic network interspersed with polymer-free liquid regions. Thus, as concentration was increased, hydrogel formation appeared to coincide with intergrowth of microscopically phase-separated domains of 1–25 μm in size. This heterogeneity suggests that the block copolypeptides preferentially segregate into gel matrix domains that exclude some fraction of the available water. CTEM visualization of the vitrified gel domains (Supplementary Information) revealed a complex membrane assembly responsible for network formation at the nanoscale. Similar morphologies were observed for all other hydrogel-forming polypeptide amphiphiles.

The porous nature of the copolypeptide gels was verified using microrheological experiments. This technique relies on the measurement of thermally induced brownian motion of micrometre-size tracer particles dispersed in the sample^{19,20}. In a 0.25-wt% $K_{160}L_{40}$ gel it was observed that tracer particles of 0.5 μm diameter diffused freely as if suspended in a liquid environment (that is, their mean-square displacement, MSD, increased linearly with time), whereas 1.0- μm particles showed restricted motion (MSD reached a plateau over time) as they began to experience some confinement. This shift in behaviour was consistent with both

the LSCM and earlier rheological data (Fig. 2a) in that the larger 1.0- μm tracers should experience some restraint as they are near the lower size limit of the voids. Microrheology confirmed the existence of microscopic heterogeneity in these hydrogels, which was found to persist after prolonged standing, heating, centrifugation, and bulk shear of the samples.

This morphology is unusual in that most hydrogels are heterogeneous at short length scales (<100 nm), but homogenous at the microscopic level^{1–4}. In the copolypeptides, the hydrophobic associations must be sufficiently strong to stop the swelling driven by electrostatic repulsion of the polyelectrolyte segments. The strength of these associations probably originates from the ordered packing of the α -helical and β -strand segments. The resulting porous structure represents a balance of these competing attractive and repulsive forces, and may prove advantageous in developing these materials via suitable engineering for biotechnology applications where diffusion of large molecules (for example, proteins, DNA) through the voids²¹, entrapment and proliferation of cells for tissue regeneration²², or the biomimetic confined nucleation of inorganic phases^{23,24} are desired.

The dynamics of the gelation process were also investigated. In the rheometer, the polymer gels were subjected to large amplitude oscillations to break down the gel structure. The recovery of the gels was then probed by measuring G' and G'' in the linear, small deformation regime as a function of time. Within the short time that had elapsed in between switching amplitudes (about 10 s), the hydrogels had recovered 80–90 % of their strength, followed by a slower reorganization process where the full initial value of G' was restored (Fig. 2b). This rapid recovery can be attributed to the physical nature of gelation and the relatively low relative molecular mass of the copolypeptides, which enables the molecules to reorganize quickly. Gelatin gels, because of the low mobility of large macromolecules, recover only slowly after being broken down by shear. The low gelation concentrations (~1.0 wt%) for these polypeptides can also be contrasted with surfactant-based hydrogels that require considerable amounts of low relative molecular mass lipid (20–50 wt%) to achieve gelation²⁵. The exceedingly low mass fraction of material in the polypeptide hydrogels, combined with their recovery properties and microporous structure, allows them to fill an advantageous and unique niche in between conventional polymer and surfactant hydrogels. Their peptide backbone also imparts to these materials advantageous features of proteins (such as degradability and functionality), which makes them attractive for biomedical applications. □

Methods

Synthesis

All block copolypeptides were synthesized using $\text{Co}(\text{PMe}_3)_4$ initiator²⁶, and were purified and then characterized using size exclusion chromatography, ^1H and ^{13}C nuclear magnetic resonance (NMR), and infrared spectroscopy according to literature procedures⁹. Isolated yields of the final copolymers ranged between 75% and 90%. Amino-acid compositions of the copolymers were found to be within 3% of predicted values. Chain lengths of the copolymers were found to be within 8% of predicted lengths with CLD (weight average length/number average length) ranging between 1.1 and 1.3 (Supplementary Information).

Circular dichroism

Copolypeptides, 0.5 mg ml⁻¹ in deionized water, were analysed in a 0.5-mm path length quartz cell on an Olis RSM 1000 Spectrometer. Samples analogous to $K_{160}L_{40}$, $K_{160}V_{40}$, $K_{180}L_{20}$, $K_{90}L_{10}$, $K_{160}(\text{rac-L})_{40}$, and $K_{180}V_{20}$ were prepared using racemic D/L-lysine to eliminate contributions of the polylysine domain from the CD spectra²⁷. The approximate conformations of the hydrophobic domains were determined using the method of Fasman²⁸. $K_{160}L_{40}$ and $K_{180}L_{20}$: 90% α -helix, 10% random coil. $K_{90}L_{10}$: 50% α -helix, 50% random coil. $K_{160}(\text{rac-L})_{40}$: no optical activity. $K_{160}V_{40}$: 80% β -sheet, 20% random coil. $K_{180}V_{20}$: 40% β -sheet, 40% α -helix, 20% random coil (note the double minimum for $K_{180}V_{20}$ in Fig. 3d indicating the presence of partial α -helical structure).

Rheology

Rheological measurements were performed on a Rheometrics ARES controlled strain rheometer in cone-plate geometry with 50 mm diameter and 2° cone angle. For each sample, small-deformation linearity was checked before performing oscillatory

measurements. All hydrogels were prepared by dissolving freeze-dried polypeptide powder in deionized water. To speed up dissolution, vortex mixing was also applied for a few minutes to the samples, although identical samples could be prepared without agitation by letting them stand overnight. Gelatin samples were prepared by dissolving the protein in warm water (60 °C).

Microrheological measurements were performed by dispersing a small amount of monodisperse polystyrene spheres (with amidine surface groups) into copolymer solutions of K₁₆₀L₄₀. The brownian motion of the tracer particles was captured by digital video microscopy in a conventional microscope (Nikon) with 100 × oil-immersion objective. Image analysis was performed with IDL software. Particle trajectories were extracted and analysed using algorithms developed and kindly provided by the Weitz group (Harvard University)¹⁹.

Microscopy

For LSCM, lipophilic fluorescent dye (DiOC₁₈, Molecular Probes) was dissolved (about 0.001 wt%) in THF. Several drops of this dye solution were added to deionized water that was used to prepare the hydrogels, which were allowed to stand for over 12 h before imaging. Imaging was performed with a Zeiss 510 microscope equipped with an ArKr laser (30 mW) using 488 nm as the excitation line for the dye. For CTEM, a 1.0-wt% sample of K₁₆₀L₄₀ was spread into a thin film on glass, into which a carbon-coated TEM grid was pressed, thus transferring a thin film of sample. The hydrated gel was then vitrified on the grid in liquid ethane with a Leica cryoplunging system. The *in situ*, vitrified gel membrane scaffolding was subsequently directly imaged without staining at 200 kV in bright-field mode using a Gatan 626 cryotransfer stage in a JEOL 2000FX microscope.

Received 22 October 2001; accepted 5 April 2002.

- Okano, T. (ed.) *Biorelated Polymers and Gels* (Academic, San Diego, 1998).
- Dagani, R. Intelligent gels. *Chem. Eng. News* **75**, 23–26–37 (1997).
- Peppas, N. A., Huang, Y., Torres-Lugo, M., Ward, J. H. & Zhang, J. Physicochemical foundations and structural design of hydrogels in medicine and biology. *Annu. Rev. Biomed. Eng.* **2**, 9–29 (2000).
- Ward, A. G. & Courts, A. (eds) *The Science and Technology of Gelatin* (Academic, London, 1977).
- Petka, W. A., Harden, J. L., McGrath, K. P., Wirtz, D. & Tirrell, D. A. Reversible hydrogels from self-assembling artificial proteins. *Science* **281**, 389–392 (1998).
- Wang, C., Stewart, R. J. & Kopeček, J. Hybrid hydrogels assembled from synthetic polymers and coiled-coil protein domains. *Nature* **397**, 417–420 (1999).
- Deming, T. J. Facile synthesis of block copolypeptides of defined architecture. *Nature* **390**, 386–389 (1997).
- Deming, T. J. Cobalt and iron initiators for the controlled polymerization of alpha-amino acid-N-carboxyanhydrides. *Macromolecules* **32**(13), 4500–4502 (1999).
- Katchalski, E. & Sela, M. Synthesis and chemical properties of poly-α-amino acids. *Adv. Protein Chem.* **13**, 243–492 (1958).
- Buitenhuis, J. & Forster, S. Block copolymer micelles: viscoelasticity and interaction potential of soft spheres. *J. Chem. Phys.* **107**(1), 262–272 (1997).
- Guenoun, P. *et al.* Polyelectrolyte micelles: self-diffusion and electron microscopy studies. *Langmuir* **16**(10), 4436–4440 (2000).
- Hamley, I. W. *et al.* From hard to soft spheres: the effect of copolymer composition on the structure of micellar cubic phases formed by diblock copolymers in aqueous solution. *Langmuir* **16**(6), 2508–2514 (2000).
- Won, Y.-Y., Davis, H. T. & Bates, F. S. Giant wormlike rubber micelles. *Science* **283**, 960–963 (1999).
- Moffitt, M., Khougaz, K. & Eisenberg, A. Micellization of ionic block copolymers. *Acc. Chem. Res.* **29**, 95–102 (1996).
- Tsitiliani, C., Iliopoulos, I. & Ducouret, G. An associative polyelectrolyte end-capped with short polystyrene chains. Synthesis and rheological behaviour. *Macromolecules* **33**(8), 2936–2943 (2000).
- Clark, A. C. & Ross-Murphy, S. B. Structural and mechanical properties of biopolymer gels. *Adv. Polym. Sci.* **83**, 57–192 (1987).
- Wang, G. M. & Ross-Murphy, S. B. Rheological characterisation of polymer gels. *Prog. Polym. Sci.* **23**(3), 533–562 (1998).
- Yu, M., Nowak, A. P., Pochan, D. P. & Deming, T. J. Methylated mono- and diethyleneglycol functionalized polylysines: nonionic, helical, water soluble polypeptides. *J. Am. Chem. Soc.* **121**, 12210–12211 (1999).
- Crocker, J. C. *et al.* Two-point microrheology of inhomogeneous soft materials. *Phys. Rev. Lett.* **85**(4), 888–891 (2000).
- Mason, T. G. & Weitz, D. A. Optical measurements of frequency-dependent linear viscoelastic moduli of complex fluids. *Phys. Rev. Lett.* **74**(7), 1250–1253 (1995).
- Liu, L., Li, P. & Asher, S. A. Entropic trapping of macromolecules by mesoscopic periodic voids in a polymer hydrogel. *Nature* **397**, 141–144 (1999).
- Lee, K. Y. & Mooney, D. J. Hydrogels for tissue engineering. *Chem. Rev.* **101**, 1869–1880 (2001).
- Falini, G., Fermani, S., Gazzano, M. & Ripamonti, A. Polymorphism and architectural crystal assembly of calcium carbonate in biologically inspired polymeric matrices. *J. Chem. Soc. Dalton* **21**, 3983–3987 (2000).
- Cha, J. N., Stucky, G. D., Morse, D. E. & Deming, T. J. Biomimetic synthesis of ordered silica structures mediated by block copolypeptides. *Nature* **403**, 289–292 (2000).
- Hoffmann, H. & Ulbricht, W. Surfactant gels. *Curr. Opin. Colloid Interf. Sci.* **1**, 726–739 (1996).
- Klein, H. F. & Karsch, H. H. Methylcobalt compounds with non-chelating ligands. 1. Methyltetraakis(trimethylphosphine) cobalt and its derivatives. *Chem. Ber.* **108**(3), 944–955 (1975).
- Kubota, S. & Fasman, G. Conformation and optical properties of poly(L-valine) in aqueous solution. “A single extended β chain.” *J. Am. Chem. Soc.* **96**, 4684–4686 (1974).
- Adler, A. J., Greenfield, N. J. & Fasman, G. D. Circular dichroism and optical rotary dispersion of proteins and polypeptides. *Methods Enzymol.* **27**, 675–735 (1973).

Supplementary Information accompanies the paper on Nature's website (<http://www.nature.com>).

Acknowledgements

This work was supported by grants from the National Science Foundation (Chemical and Transport Systems, and MRSEC Program). V.B. thanks the Netherlands Organization for Scientific Research (NWO) for a Talent-grant. We thank J. Hu for assistance with NMR experiments.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to T.J.D. (e-mail: tdeming@mrl.ucsb.edu).

Low slip rates and long-term preservation of geomorphic features in Central Asia

Ralf Hetzel^{*†‡}, Samuel Niedermann^{*}, Mingxin Tao[‡], Peter W. Kubik[§], Susan Ivy-Ochs^{||}, Bo Gao[‡] & Manfred R. Strecker[†]

^{*} GeoForschungsZentrum Potsdam, Telegrafenberg, D-14473 Potsdam, Germany

[†] Institut für Geowissenschaften, Universität Potsdam, D-14415 Potsdam, Germany

[‡] The State Key Laboratory of Gas-geochemistry, Lanzhou Institute of Geology, Chinese Academy of Science, 324 Donggang West Avenue, Lanzhou 730000, China

[§] Paul Scherrer Institut, c/o the Institute of Particle Physics; and the ^{||} Institute of Particle Physics, ETH Hönggerberg, CH-8093 Zurich, Switzerland

In order to understand the dynamics of the India–Asia collision zone, it is important to know the strain distribution in Central Asia, whose determination relies on the slip rates for active faults^{1–5}. Many previous slip-rate estimates of faults in Central Asia were based on the assumption that offset landforms are younger than the Last Glacial Maximum (~20 kyr ago)^{6–11}. In contrast, here we present surface exposure ages of 40 to 170 kyr, obtained using cosmogenic nuclide dating, for a series of terraces near a thrust at the northern margin of the Tibetan Plateau. Combined with the tectonic offset, the ages imply a long-term slip rate of only about 0.35 mm yr⁻¹ for the active thrust, an order of magnitude lower than rates obtained from the assumption that the terraces formed after the Last Glacial Maximum. Our data demonstrate that the preservation potential of geomorphic features in Central Asia is higher than commonly assumed.

The ongoing collision between India and Eurasia has caused widespread Cenozoic deformation resulting in uplift of the Himalayas, the Tibetan Plateau and other mountain ranges in Central Asia¹² (Fig. 1a). The present rate of convergence between India and Eurasia is of the order of 40–50 mm yr⁻¹ (see ref. 13), with only 18 ± 2 mm yr⁻¹ being accommodated by Himalayan thrusts¹⁴. The remaining 20–30 mm yr⁻¹ of convergence are distributed on faults in actively growing mountain ranges farther north. Ideally, slip rates of active faults are determined by dividing measured tectonic offsets by well constrained ages for the faulted geomorphic markers, as, for example, has been achieved along the Kunlun fault¹⁵. However, many slip rates in Central Asia have been based on offset geomorphic features which were assumed to have formed after the Last Glacial Maximum (~20 kyr ago)^{6–11}. In Central Asia the maximum advance of glaciers apparently occurred well before 20 kyr ago, so tectonically offset landforms may in fact be much older¹⁶. This seems to be the case for the central Altyn Tagh fault at the northern boundary of the Tibetan Plateau, where inferred ages of faulted landforms result in a high slip rate of 20–30 mm yr⁻¹ (ref. 6), while

palaeoseismological investigations suggest less than 20 mm yr^{-1} slip¹⁷ and global positioning system (GPS) measurements indicate a rate of $9 \pm 5 \text{ mm yr}^{-1}$ (ref. 18).

One region where a considerable portion of the India–Eurasia convergence rate is absorbed is the more than 5,000-m-high Qilian Shan, where ongoing deformation leads to the formation of mountain ranges and intramontane basins that are successively incorporated into the laterally growing Tibetan Plateau (Fig. 1a) (see ref. 19 for a review). The Qilian Shan mountain fronts are paralleled by large, coalesced alluvial fan systems, and Quaternary period deformation advancing north-northeast toward the foreland profoundly influences landscape development^{6–11,19} (Fig. 1b). When new thrusts break the surface, sedimentation on the alluvial fan surfaces stops owing to hanging-wall uplift, and rivers crossing the faults begin downcutting and form terraces. Until now, slip rates in the Qilian Shan have been derived for only two thrusts that tectonically offset late Quaternary alluvial fans^{8,11} (Fig. 1b). At the front of the Yumu Shan, a large fore-mountain of the Qilian Shan, two spatially separated fault scarps located at an elevation of approximately 1,800 m may be part of a single south-dipping thrust fault⁸. The smaller scarp, 1.5 to 4.4 m in height, has been incised by presently active braided channels. From a series of ten scarp profiles⁸ the presence of at least two different terrace levels can be inferred.

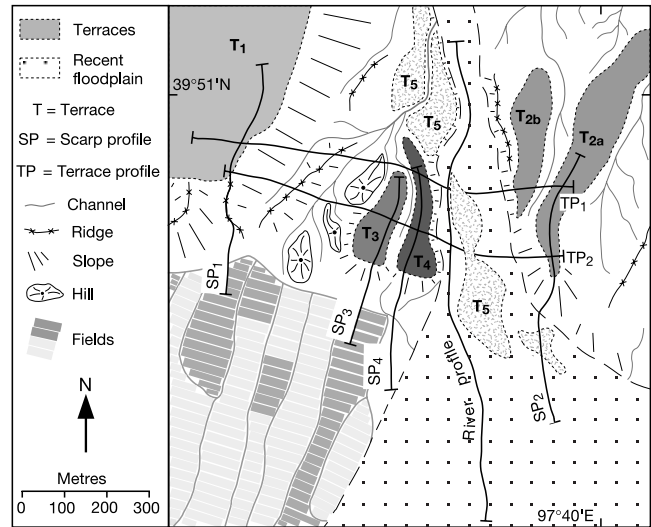


Figure 2 Map showing the spatial distribution of terraces at site 2 and the locations of four fault scarp (SP₁, SP₂, SP₃ and SP₄) and two terrace (TP₁ and TP₂) profiles. Location of map is shown in Fig. 1c.

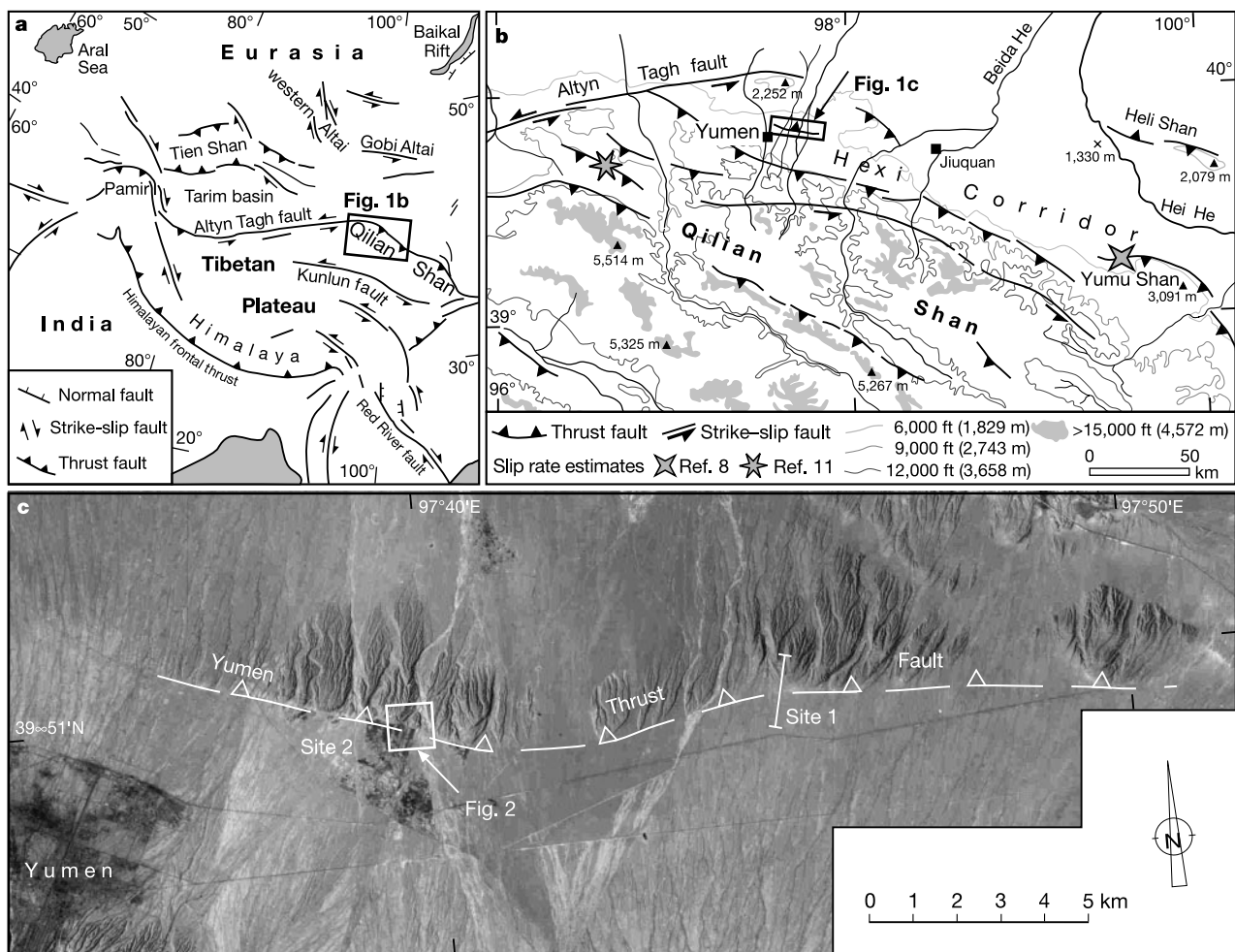


Figure 1 Geological setting of the studied thrust fault. **a**, Map of Central Asia¹² with the major active strike-slip and reverse faults and the location of **b**. **b**, Morphotectonic map of the northwestern Qilian Shan¹¹ showing the location of **c**. Grey stars indicate two reverse faults for which slip rates were proposed previously. A shortening rate of $2.1 \pm 0.6 \text{ mm yr}^{-1}$ for a thrust in the western Qilian Shan is based on the assumption that a tectonically offset alluvial fan is $11 \pm 3 \text{ kyr}$ old¹¹. In the Yumu Shan, maximum and

minimum rates of vertical uplift of 1.9 and 0.4 mm yr^{-1} were derived by assigning ages of 8 and 10 kyr to two fault scarps 3-m-high and 19-m-high, respectively⁸. **c**, Enlarged portion of a CORONA satellite image (9044 D 06 088), showing the Yumen thrust and the two study sites. Spatial resolution is about 8 m (25 feet). Active north-flowing rivers have incised into the uplifted hanging wall of the thrust and have formed terraces.

The larger 18- to 20-m-high scarp offsets an older terrace surface formed by a veneer of late Quaternary fanglomerates that discordantly overlie Pliocene epoch or older sediments⁸. Further west, close to the Altyn Tagh fault, an approximately 20-km-long thrust fault scarp at an altitude of about 3,000 m displaces fluvial or fluvio-glacial fanglomerates of a large alluvial fan¹¹. Profiles across the scarp revealed different terrace levels located 12 to 25 m above the active channels that cross the scarp. These surfaces were interpreted to reflect the composite, locally diachronic nature of the fan surface¹¹.

Here we report data from similar, seemingly youthful terraces aligned along the active Yumen thrust fault in the western Qilian Shan¹¹. This thrust is located about 10 km north of the Qilian Shan mountain front at an elevation of 2,000 m (Fig. 1c). The north-dipping fault has displaced the surface of an alluvial fan and forms a 25-km-long south facing scarp. North-flowing rivers draining the Qilian Shan have incised into the uplifting hanging wall and formed several terrace levels. The vertical displacement of the thrust and the elevations of the terraces above the active river channels were determined by high-precision topographic profiling in the central (site 1) and western sectors (site 2) of the thrust (Fig. 1c). At site 2, where an active river has formed six terraces (Fig. 2), two east–west profiles constrain the vertical distance between the individual terraces and the active river channel (Fig. 3a). The displacement on the fault since formation of the terraces is derived from five north–south scarp profiles (Fig. 3b). The small difference between the displacement extracted from the scarp profiles and the height of the terraces above the river is caused by minor volumes of sediment that were deposited in front of the scarp by the river. Thus, although three of the scarp profiles at site 2 slightly underestimate the vertical slip since terrace formation, the distance between the river and the

terraces should reflect the correct vertical displacements for all terraces. The highest identified terrace level is located at site 1, where an active river channel and the base of the fault scarp are at the same elevation, and a single scarp profile reveals a displacement of 57 m (Fig. 3b). The decreasing height of the fault scarp from the centre to the western part of the thrust most probably results from lateral propagation of the Yumen fault.

Terrace ages were determined by ^{21}Ne , ^{10}Be and ^{26}Al surface exposure dating using quartz clasts. A description of the terraces, the sampling strategy, the analytical procedures and the applied correction for the inherited cosmogenic nuclide component is given below (see Methods). The surface exposure ages obtained for the different terraces are high (Fig. 4 and Supplementary Information). The oldest terrace located at site 1 has an age of ~ 170 kyr, as constrained by identical ^{21}Ne , ^{10}Be and ^{26}Al ages. The fact that the different cosmogenic nuclides yield ages that agree within limits of error, as well as the consistency of ages with terrace stratigraphy, clearly documents that all terraces have been dated reliably and are considerably older than the Last Glacial Maximum. We note that the apparent inconsistency of the ages for terraces T_1 and T_{2a} with terrace stratigraphy is insignificant because within error limits terrace T_1 may still be older than terrace T_{2a} .

The position of the terraces in the hanging wall of the Yumen thrust suggests that tectonic uplift exerted the largest control on terrace formation. Assuming that the slip rate of the thrust has remained constant during its entire evolution and that river incision has kept pace with hanging wall uplift, the terrace ages should define a linear trend when plotted versus the heights of the terraces above the active channels (Fig. 4). To a first approximation, the data are in accordance with such an interpretation, because nearly all ages fall between two linear trend lines defined by river incision rates of 0.3

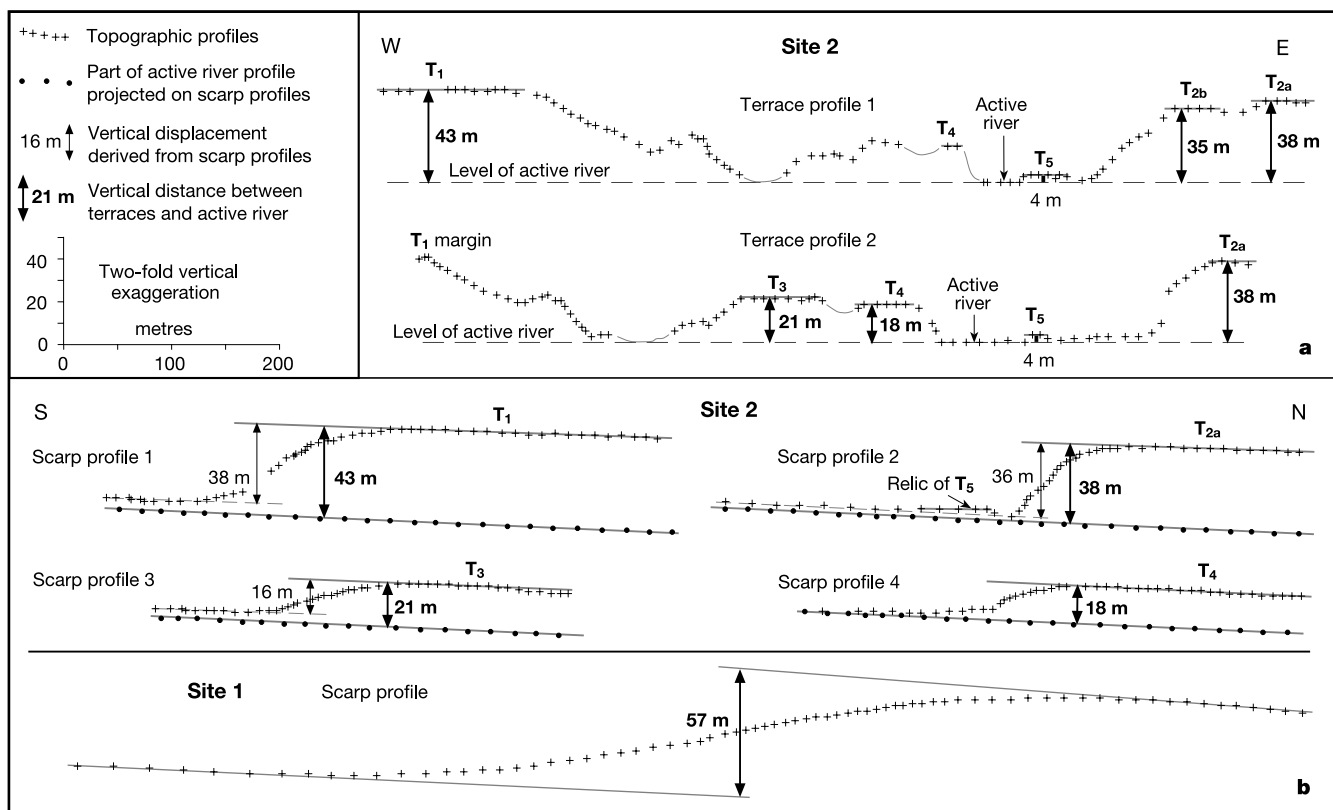


Figure 3 Topographic profiles at the Yumen thrust. All levelling points have been projected from the profile lines shown in Fig. 2 onto vertical planes. Vertical exaggeration is two-fold for all profiles. **a**, Two east–west terrace profiles at site 2 showing the height of

the terraces above the active river channel. **b**, Five north–south fault scarp profiles extending from the different terraces down the fault scarp. The levelling points from a river profile (see Fig. 2) have been projected onto the four scarp profiles at site 2.

and 0.4 mm yr^{-1} . Thus, we suggest that the average long-term vertical slip rate of the Yumen thrust is $0.35 \pm 0.05 \text{ mm yr}^{-1}$. As a consequence, rates of faulting in the Qilian Shan, hitherto assumed to be of the order of a few millimetres per year, may be considerably smaller. The age of the lowermost terrace, T_5 , which is located only 4 m above the active channel and has not been sampled for dating, can be estimated by combining the inferred long-term slip rate with the elevation of the terrace. This suggests that terrace T_5 is 10 to 14 kyr old.

Apart from tectonic uplift, the incision rate of rivers and the formation of terraces may be significantly influenced by variations in climate (for example, see refs 20, 21) and this may explain the deviation of exposure ages from the linear trend of about 0.35 mm yr^{-1} . Discharge and sediment load of the perennial streams draining the Qilian Shan were presumably variable during the course of frequent Quaternary climatic changes. Periods of warming and deglaciation have most probably increased the surface runoff and the rivers' ability to erode and downcut, whereas during cold periods river incision slowed down. As a consequence, we expect that the superposition of a variable climate signal onto a constant tectonic rate should result in river-incision rates that oscillate around an average rate imposed by faulting. In Fig. 4, which relates terrace age and corresponding elevation, periods of relatively high or low incision rates correlate with steeper or shallower slopes, respectively. When correlating terrace ages with palaeoclimatological records that have a higher temporal resolution it should be kept in mind that systematic errors of cosmogenic nuclide production rates may shift absolute ages by around 10% in either direction but will not affect relative age differences. The formation of terraces T_1 and T_{2a} seems to be related to a relatively high incision rate at 110 to 120 kyr ago and may correlate with a warm period immediately following the penultimate deglaciation (see ref. 22 for example). Another period of high incision has occurred around 80 kyr ago and probably reflects relatively humid conditions at the end of the last interglacial period. In contrast, a low incision rate between 60 and 40 kyr ago may reflect a cold

period at the beginning of the last glaciation. Renewed incision after formation of terrace T_4 correlates well with a moist period of warming between 40 and 32 kyr ago documented by the $\delta^{18}\text{O}$ and anion-concentration records in an ice core from the Qilian Shan²³. Thus the terraces at the Yumen thrust record both a presumably constant tectonic uplift rate and variable climate conditions.

In conclusion, surface exposure dating and topographic surveying of terraces at the Yumen thrust in the Qilian Shan demonstrate unequivocally that: (1) parts of the present-day landscape are considerably older than the ultimate transition from glacial to present interglacial conditions; (2) previous slip-rate estimates for many faults in Central Asia may be biased toward high values; (3) cosmogenic nuclide dating allows us to determine accurate rates of faulting for periods of about 10^5 years; and (4) landscapes in Central Asia make it possible to study the relationship between tectonics, surface processes and superimposed climate fluctuations during at least the past 200 kyr. □

Methods

Terrace description, sampling strategy and sample preparation

The terraces at the Yumen thrust bear no vegetation and are devoid of a silt cover. Clasts on the terrace surfaces range in size from 0.5 to 5 cm and are in contact with each other. Minor amounts of sand and silt are restricted to the interstitial space between the clasts. We collected 100 to 300 quartz clasts from the most pristine part of each terrace. Terrace microtopography is only on the order of a few centimetres, suggesting negligible erosion. Sampling concentrated on quartz clasts for the following reasons. Granite and gneiss clasts are rare and it is impossible to collect a sufficient number of clasts to apply the amalgamation approach necessary to correct for the inherited cosmogenic nuclide component (see below). On the other hand, abundant phyllite clasts were not sampled because (1) separation of quartz from this fine-grained rock-type is difficult and (2) phyllites may have high concentrations of U and Th and thus may contain nucleogenic Ne, which is produced by (α, n) reactions from ^{18}O and ^{19}F and may render the distinction of cosmogenic Ne difficult. For each sample, more than 30 of the quartz clasts collected in the field were chosen. The selected clasts, which had a similar size of 2–3 cm, were crushed and sieved in different size fractions. The chemical treatment involved a first leach in HCl at 80°C , followed by a series of three leaches in a dilute HF/HNO_3 mixture at 80°C in an ultrasonic bath²⁴. The etching in the HF/HNO_3 mixture effectively removed all meteoric ^{10}Be , as shown by Al concentrations of only 25–30 p.p.m. in the quartz separates used for ^{26}Al exposure dating.

Analytical procedures

Noble gas analysis of the quartz separates was carried out in the noble gas laboratory of the GFZ Potsdam on a VG5400 mass spectrometer. Gas extraction was accomplished by stepwise heating at 400, 600, 800 and $1,700^\circ\text{C}$. In addition, aliquots of all samples were crushed in vacuum, as in the course of this work it turned out that the isotopic composition of Ne trapped in fluid inclusions differs from that of atmospheric Ne²⁵. Therefore, to determine the concentration of cosmogenic ^{21}Ne , the ^{21}Ne excess in the step-heated samples was calculated relative to the trapped component as revealed by crushing. Consistent ^{21}Ne , ^{10}Be and ^{26}Al ages confirm the validity of our approach (see Fig. 4 and the Supplementary Information). Assuming an atmospheric composition of the trapped Ne, instead of the component revealed by crushing, would result in erroneously high ^{21}Ne ages for several samples²⁵. Aliquots of some quartz samples were analysed for ^{10}Be and ^{26}Al using accelerator mass spectrometry (AMS) at the PSI/ETH facility in Zurich. The chemical separation of Be and Al and the AMS measurements were carried out as described²⁶.

Correction for inheritance and age calculation

When dating surfaces such as terraces it is important to take into account that cosmogenic nuclides are not only produced after deposition of a clast but are already produced during erosion of the host rock and sedimentary transport of clasts. To correct for this inherited nuclide component, we used the amalgamation approach^{27,28}. We have taken amalgamated samples at 0.6, 1.2 and 1.8 m depth from a hole excavated on terrace T_1 and from two active river channels, one in the central part of the thrust near site 1 and the other one from the river that formed the terraces at site 2. The amount of inherited cosmogenic nuclides as derived from Ne and Be analyses of the subsurface samples and the active river sample at site 2 agrees well and corresponds to an age of ~ 10 kyr. As nuclide inheritance may be variable for different terraces^{27,28} we have enlarged the error limits of the inheritance correction applied to site 2 to 10 ± 6 kyr. At site 1, the inherited component corresponds to an age of 30 ± 5 kyr, as revealed by Be and Al analyses of the active river sample. Increasing the uncertainty for the reason outlined above, we use 30 ± 9 kyr as an inheritance correction. The distinct values for the inherited component at the two sites are readily explained by the different sizes and mean elevations of the catchment areas of the two rivers that cross the Yumen thrust²⁵. To calculate exposure ages from cosmogenic nuclide concentrations corrected for the inherited component, the production rates of ^{21}Ne , ^{10}Be and ^{26}Al (refs 26, 29) were scaled to the sampling locations applying the method described in ref. 30. All ages have been calculated assuming no erosion. Errors on the

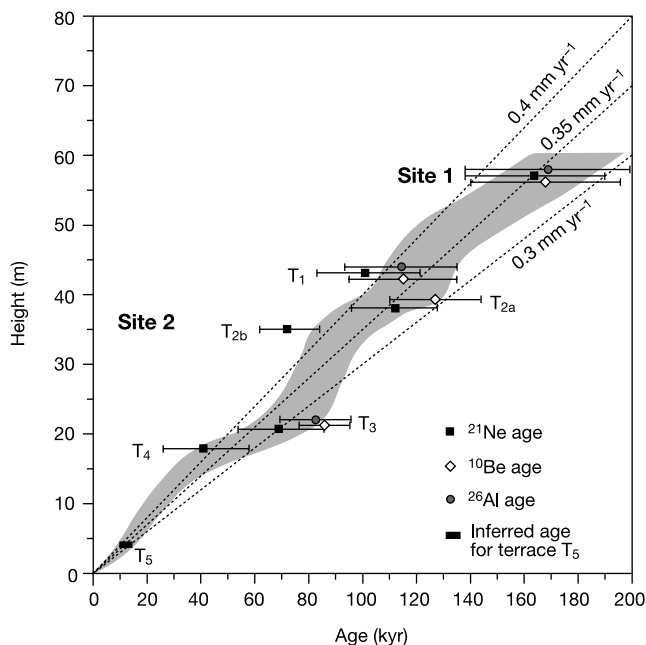


Figure 4 Plot of exposure ages versus heights of the different terraces. Error bars indicate 2σ values. Three dashed lines indicate constant rates of river incision and tectonic uplift of 0.3, 0.35, and 0.4 mm yr^{-1} . The shaded curve illustrates qualitatively that incision rates have temporally changed as a result of climate variations superposed on a constant rate of faulting.

exposure ages do not include the systematic uncertainties of the production rates, which are of the order of 10%.

Received 16 January; accepted 16 April 2002.

1. Avouac, J.-P. & Tapponnier, P. Kinematic model of active deformation in central Asia. *Geophys. Res. Lett.* **20**, 895–898 (1993).
2. Peltzer, G. & Saucier, F. Present-day kinematics of Asia derived from geologic fault rates. *J. Geophys. Res.* **101**, 27943–27956 (1996).
3. England, P. & Molnar, P. The field of crustal velocity in Asia calculated from Quaternary rates of slip on faults. *Geophys. J. Int.* **130**, 551–582 (1997).
4. Holt, W. E. Correlated crust and mantle strain fields in Tibet. *Geology* **28**, 67–70 (2000).
5. England, P. & Molnar, P. Active deformation of Asia: from kinematics to dynamics. *Science* **278**, 647–650 (1997).
6. Peltzer, G., Tapponnier, P. & Armijo, R. Magnitude of Late Quaternary left-lateral displacement along the north edge of Tibet. *Science* **246**, 1285–1289 (1989).
7. Peltzer, G. *et al.* Offsets of late Quaternary morphology, rate of slip, and recurrence of large earthquakes on the Chang Ma fault (Gansu, China). *J. Geophys. Res.* **93**, 7793–7812 (1988).
8. Tapponnier, P. *et al.* Active thrusting and folding in the Qilian Shan, and decoupling between upper crust and mantle in northeastern Tibet. *Earth Planet. Sci. Lett.* **97**, 382–403 (1990).
9. Gaudemer, Y. *et al.* Partitioning of crustal slip between linked active faults in the eastern Qilian Shan, and evidence for a major seismic gap, the 'Tianzhu Gap', on the western Haiyuan fault, Gansu (China). *Geophys. J. Int.* **120**, 599–645 (1995).
10. Meyer, B. *et al.* Rate of left-lateral movement along the easternmost segment of the Altyn Tagh fault, east of 96°E (China). *Geophys. J. Int.* **124**, 29–44 (1996).
11. Meyer, B. *et al.* Crustal thickening in Gansu-Qinghai, lithospheric mantle subduction, and oblique, strike-slip controlled growth of the Tibet plateau. *Geophys. J. Int.* **135**, 1–47 (1998).
12. Molnar, P. & Tapponnier, P. Cenozoic tectonics of Asia: effects of a continental collision. *Science* **189**, 419–426 (1975).
13. DeMets, C., Gordon, R. G., Argus, D. F. & Stein, S. Effect of recent revisions to the geomagnetic reversal time scale on estimates of current plate motions. *Geophys. Res. Lett.* **21**, 2191–2194 (1994).
14. Project Idylhim members, Bilham, R., Larson, K. & Freymueller, J. GPS measurements of present-day convergence across the Nepal Himalaya. *Nature* **386**, 61–64 (1997).
15. Van Der Woerd, J. *et al.* Uniform postglacial slip-rate along the central 600 km of the Kunlun Fault (Tibet), from ²⁶Al, ¹⁰Be, and ¹⁴C dating of riser offsets, and climatic origin of the regional morphology. *Geophys. J. Int.* **148**, 356–388 (2002).
16. Gillespie, A. & Molnar, P. Asynchronous maximum advances of mountain and continental glaciers. *Rev. Geophys.* **33**, 311–364 (1995).
17. Washburn, Z. *et al.* Late Holocene earthquake history of the central Altyn Tagh fault, China. *Geology* **29**, 1051–1054 (2001).
18. Bendick, R., Bilham, R., Freymueller, J., Larson, K. & Yin, G. Geodetic evidence for a low slip rate in the Altyn Tagh fault system. *Nature* **404**, 69–72 (2000).
19. Tapponnier, P. *et al.* Oblique stepwise rise and growth of the Tibetan Plateau. *Science* **294**, 1671–1677 (2001).
20. Bull, W. B. Threshold of critical power in streams. *Geol. Soc. Am. Bull.* **90**, 453–464 (1979).
21. Molnar, P. *et al.* Quaternary climate change and the formation of river terraces across growing anticlines on the north flank of the Tien Shan, China. *J. Geol.* **102**, 583–602 (1994).
22. Imbrie, J. *et al.* in *Milankovitch and Climate, Part 1* (eds Berger, A. L. *et al.*) 269–305 (Reidel, Boston, 1984).
23. Thompson, L. G. *et al.* Holocene-late Pleistocene climatic ice core records from the Qinghai-Tibetan plateau. *Science* **246**, 474–477 (1989).
24. Kohl, C. P. & Nishiizumi, K. Chemical isolation of quartz for measurement of *in-situ*-produced cosmogenic nuclides. *Geochim. Cosmochim. Acta* **56**, 3583–3587 (1992).
25. Hetzel, R. *et al.* ²¹Ne versus ¹⁰Be and ²⁶Al exposure ages of fluvial terraces: The influence of crustal Ne in quartz. *Earth Planet. Sci. Lett.* (submitted).
26. Kubik, P. W., Ivy-Ochs, S., Masarik, J., Frank, M. & Schlüchter, C. ¹⁰Be and ²⁶Al production rates deduced from an instantaneous event within the dendro-calibration curve, the landslide of Kofels, Ötztal Valley, Austria. *Earth Planet. Sci. Lett.* **161**, 231–241 (1998).
27. Repka, J. L., Anderson, R. S. & Finkel, R. C. Cosmogenic dating of fluvial terraces, Fremont River, Utah. *Earth Planet. Sci. Lett.* **152**, 59–73 (1997).
28. Hancock, G. S., Anderson, R. S., Chadwick, O. A. & Finkel, R. C. Dating fluvial terraces with ¹⁰Be and ²⁶Al profiles: application to the Wind River, Wyoming. *Geomorphology* **27**, 41–60 (1998).
29. Niedermann, S. The ²¹Ne production rate in quartz revisited. *Earth Planet. Sci. Lett.* **183**, 361–364 (2000).
30. Dunai, T. J. Scaling factors for production rates of *in situ* produced cosmogenic nuclides: a critical reevaluation. *Earth Planet. Sci. Lett.* **176**, 157–169 (2000).

Supplementary Information accompanies the paper on Nature's website (<http://www.nature.com>).

Acknowledgements

R.H. acknowledges logistic support from the Chinese Academy of Sciences in Lanzhou and is grateful to J. Erzinger for his support. We thank B.P. Wernicke and G.S. Hancock for constructive reviews. This project is financially supported by the Deutsche Forschungsgemeinschaft, the Chinese Academy of Science ('Knowledge & Creativeness Project'), and the GeoForschungsZentrum Potsdam.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to R.H. (e-mail: hetzel@gfz-potsdam.de).

Recovery of 16S ribosomal RNA gene fragments from ancient halite

Steven A. Fish*, Thomas J. Shepherd†, Terry J. McGenity‡ & William D. Grant*

* Department of Microbiology and Immunology, University of Leicester, Leicester LE1 9HN, UK

† British Geological Survey, Keyworth, Nottingham NG12 5GG, UK

‡ Department of Biological Sciences, University of Essex, Colchester, Essex CO4 3SQ, UK

During the last decade, sensitive techniques for detecting DNA have been successfully applied to archaeological and other samples that were a few hundred to a few thousand years old^{1,2}. Nevertheless, there is still controversy and doubt over claims of exceptionally ancient DNA³. Additional accounts stretching back nearly a century suggest that microorganisms may survive over geological time in evaporite deposits^{4,5}. There is, however, often doubt over the age relationship between evaporite formation and the incorporation of microorganisms⁶. Here, we have used petrographic and geochemical techniques (laser ablation microprobe-inductively coupled plasma-mass spectrometry) to verify the estimated geological age of halite (NaCl) evaporite samples. Fragments of 16S ribosomal RNA genes were detected by polymerase chain reaction amplification of DNA extracted from halite samples ranging in age from 11 to 425 Myr (millions of years). Haloarchaeal 16S rDNA amplicons were present in one sample (11–16 Myr), whereas other samples (65–425 Myr) yielded only bacterial 16S rDNA amplicons. Terminal restriction fragment length polymorphism analyses indicate complex and different populations of microorganisms or their free DNA in ancient halites of different ages.

Evaporite deposits, usually halite 'rock salt' (NaCl), represent the desiccated remains of ancient seas or saline lakes. Accounts stretching back over almost a century suggest that microorganisms may survive in halite, preserved and isolated from the present-day environment⁴. Experiments using laboratory-grown halite crystals have demonstrated that microorganisms, in particular haloarchaea, become trapped in microscopic droplets of brine within the solid crystal where they remain active over several months⁷. Viable haloarchaea can still be recovered from crystals grown in 1988 (W.D.G., unpublished results). These observations provide a conceptual framework to understand and assess repeated claims that microorganisms can be isolated from ancient halite. However, during geological time, evaporites undergo varying degrees of fracturing and recrystallization. During these events, anthropogenic and geogenic processes could contaminate the salt with microorganisms of recent origin or of indeterminate age. 16S rRNA gene sequence analysis of isolates from surface hypersaline environments and ancient evaporites have failed to show consistent differences in gene sequence between isolates from different geological formations^{4,5,8}.

This study has made use of advanced geochemical and microbiological technologies to isolate and recover traces of DNA preserved in halite. Most previous studies have used large pieces of salt (up to 500 g) containing a mixture of halite and minor clay minerals, for which it is impossible to say with certainty that there has been no secondary recrystallization and/or introduction of externally derived DNA. Furthermore, any microorganisms or DNA isolated from such heterogeneous samples could have resided between rather than inside primary crystals. To minimize these uncertainties, we have used as our target halite, physically isolated single-crystal, brine-inclusion-rich microsamples. By using laser ablation microprobe-inductively coupled plasma-mass spec-

Table 1 Halite samples selected for DNA recovery

Sample type	Age of salt	Geographical location and geological setting	Type and size of brine inclusions
Poland no. 4	Middle Miocene (Badenian) 11–15.8 Myr	Wieliczka mine, Poland. Bedded evaporites, 'green salt' bed, Forecarpathian basin.	Cuboid, liquid-only, up to 150 μm^3 in chevron halite
Thailand no. 541	Upper Cretaceous 65–96 Myr	Bore hole, depth 381.5 m. Khorat plateau, northeast Thailand. Bedded evaporites, Maha Sarakham formation.	Cuboid, liquid-only, up to 120 μm^3 in chevron halite
Brazil no. 1	Lower Cretaceous (Aptian) 113–119 Myr	Bore hole, depth 768.16 m. Sergipe, Brazil. Bedded evaporites, Ibura member, Muribeca formation, Sergipe basin.	Cuboid, liquid-only, up to 100 μm^3 in chevron halite
Michigan no. 1316	Upper Silurian (Cayugan) 415–425 Myr	Bore hole, Wyandotte Chemical Corp., Michigan, USA. Bedded evaporites, Salina group, Michigan basin.	Cuboid, liquid-only, up to 150 μm^3 in chevron halite
Michigan no. 1506	Upper Silurian (Cayugan) 415–425 Myr	Bore hole, MCGC Willow Cavern, Michigan, USA. Bedded evaporites, Salina group, Michigan basin.	Cuboid, liquid-only, up to 100 μm^3 in chevron halite
Michigan state	Upper Silurian (Cayugan) 415–425 Myr	Bore hole, State USA 1-3, Michigan, USA. Bedded evaporites, A1 formation, Salina group, Michigan basin.	Cuboid, liquid-only, up to 80 μm^3 in chevron halite

trometry (LAMP–ICP–MS), we have further established that the brine inclusions are chemically representative of the original mother liquor.

Only specific types of halite satisfying the following criteria were selected for detailed validation (Table 1): bedded salt showing primary halite crystal accumulation textures; primary halite with abundant 'chevron' crystals (crystals that have grown in open brine pools)⁹; chevron crystals with dense arrays of primary brine inclusions decorating growth zones parallel to the cube faces (Fig. 1); crystals free from open microfractures or planes of secondary brine inclusions trapped along healed microfractures; monophase, liquid-only inclusions (that is, indicative of low formation temperatures). The zoning, as shown in Fig. 1, is considered a primary growth fabric and has been widely reported in natural and laboratory-grown halite crystals¹⁰. During the early stages of diagenesis, substantial recrystallization takes place; however, in the case of chevron halite, this process is only partial and relic chevron crystals are preserved. To our knowledge, no reports have been published of chevron textures having formed as a result of secondary recrystallization during or after burial. It is our belief therefore that the brine inclusions were trapped at the time of salt formation, the minimum age being fixed by the age of the immediately overlying sediments. Absolute radiometric ages for the salts are not available. For those samples that satisfied the above criteria, three were selected for special treatment (Poland no. 4, Thailand no. 541 and Michigan

state). For these, referred to as 'quality controlled' material, small portions containing high densities of brine inclusions were excised using a 532 nm laser as a surgical scalpel.

Chemical data for the brine inclusions were obtained at two independent laboratories (British Geological Survey, Nottingham and Eidgenossche Technische Hochschule (ETH) Zurich) using well documented LAMP–ICP–MS procedures^{11–13}. Although it is recognized that the chemical composition of evaporated brines may vary owing to the mixing of marine and non-marine waters, our data strongly support a seawater origin for the inclusion brines. Weight ratios for conservative elements such as K and Mg are consistent with the deposition of halite from naturally evaporated sea water. Mg/K ratios range from 6.0 to 0.2; the highest value being obtained for the Cretaceous period sample (Brazil no. 1). The latter sample also displays high Ca/Mg ratios, suggesting that the parent sea water evolved along a CaCl_2 -enrichment trend^{14,15}. The remaining five samples show a characteristic Mg-depletion with respect to modern sea water (Mg/K 0.2–4), a feature noted by other workers¹⁶. Sample Thailand no. 541 was also analysed by cryo-environmental scanning electron microscopy–energy dispersive X-ray spectroscopy (ESEM–EDS) (T. Lowenstein, personal communication)^{13,17} and yields a Mg/K mole ratio of 2.9, which lies very close to the Harvie–Moller–

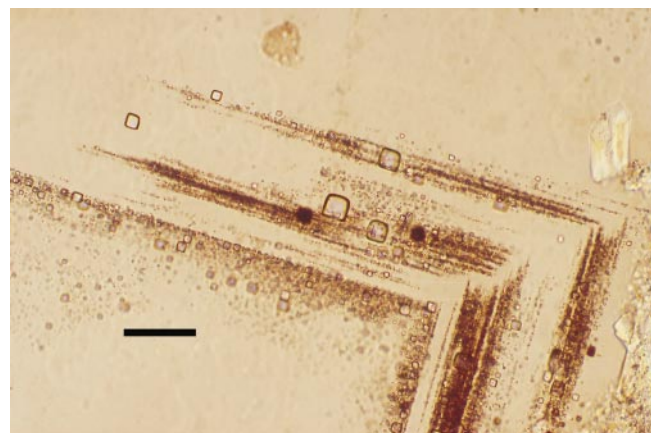


Figure 1 Photomicrograph of a typical chevron halite (taken in transmitted light) showing well developed growth zones parallel to the cube faces decorated with thousands of microscopic, cuboid brine inclusions. In this sample the primary brine inclusions range from >40 to <1 μm in size (halite from Upper Miocene salt deposits, Lorca, Spain). Scale bar, 100 μm .

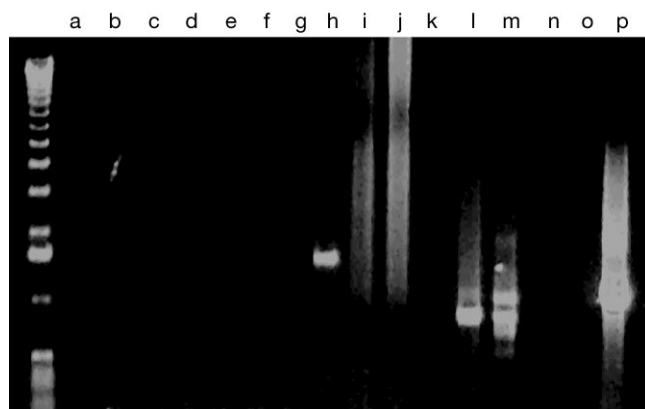


Figure 2 Sizing and visualization by agarose gel electrophoresis of amplified DNA from nested PCR using primers 27Fa/1524Ra in the first round of amplification (lanes a–h) and 353Fa/1392Ru in the second round (lanes i–p). Extractions of single chips of quality controlled Poland no. 4, Thailand no. 541 and Michigan State halite (first round lanes a, b, c and second round i, j, k, respectively) were compared to extractions derived from 3–4 pooled chips of the same material (first round lanes d, e, f and second round lanes l, m, n, respectively). Negative controls for rounds 1 and 2 were in lanes g and o, with positive controls in lanes h and p. The first lane shows size markers (*HindIII* Digest).

Weare (HMW) computer simulated modern seawater evaporation curve¹⁸. By comparison, Poland no. 4 is chemically anomalous. This sample shows strong depletion in Mg and K, suggesting an influx of non-marine brines and/or the syn-depositional recycling of pre-existing salts¹⁸. Nevertheless, the trapped inclusions are primary. During halite precipitation, the Cl/Br ratio of the parent brine decreases linearly from around 292 (value for modern sea water¹⁹) to less than 60, the point at which potassium salts begin to precipitate. Measured Cl/Br ratios for the brine inclusions (excluding Poland no. 4) are consistently less than 292. If the inclusions had been trapped in response to younger hydrogeological processes (or historical events), they would have exhibited depletions in K, Mg and Br (with respect to Na and Cl), and displayed significantly different elemental ratios (perhaps by a factor of 10). Together with the preservation of unequivocal growth textures, the weight of evidence suggests that the brine inclusions are primary and hence equivalent in age to the age of the halite.

DNA was extracted from quality controlled material and halite that was not quality controlled (that is, contained visible open and/or healed microfractures). For the material that was not quality controlled, 16S rDNA amplification products were obtained from Poland no. 4, Thailand no. 541 and Michigan no. 1316. 16S rDNA amplification products were not obtained for Brazil no. 1 and Michigan no. 1506 (data not shown). For the quality controlled material, amplified DNA was obtained for Poland no. 4, and Thailand no. 541; the Michigan State sample was negative (Fig. 2).

The positive control bands (lanes h and p of Fig. 2) increase in intensity and decrease in size from round 1 (approximate size 1,500 base pairs (bp), lane h) to round 2 (approximate size 1,000 bp, lane p), indicative of the nested PCR method used in this study. Negative controls (lanes g and o), which consisted of PCR components only, indicate that no contaminant was carried within the reaction mixture. Positive signals from the Poland and Thailand samples are seen in lanes l and m. Multiple bands (lanes l and m, approximate sizes 700–1,000 bp) suggest the presence of truncated products, possibly due to nonspecific primer annealing²⁰. There did seem to be a template dose effect in that material from single chips

failed to yield discrete bands after PCR amplification (lanes i and j), whereas material extracted from 3–4 pooled chips produced amplicons of specific sizes (lanes l and m). The Michigan State sample (lane n) gave no amplification product, supporting the view that the DNA extraction procedure was free from contamination.

Terminal restriction fragment length polymorphism (t-RFLP) analysis was used to determine the microbial diversity in and between Poland and Thailand quality controlled halite samples. The peaks on the profiles (Fig. 3) represent terminal fragments of a specific size. Six peaks were detected in the Poland sample and eleven in the Thailand sample when digested with *MspI*. Eight peaks were identified in the Poland sample and nine in the Thailand sample when digested with *RsaI*. The peak numbers represent the minimum microbial species diversity in each sample. The *MspI* digests shared three peaks and the *RsaI* digests shared two. These profiles clearly indicate that the microbial DNAs trapped within the Poland and Thailand samples are significantly different.

The amplified 16S rRNA genes from Polish halite (quality controlled and not quality controlled), Thailand halite (quality controlled) and Michigan no. 1316 halite (not quality controlled) were cloned, sequenced and subjected to phylogenetic analysis (Fig. 4a, b). In view of the archaeal-specific primers used and the repeated isolation of haloarchaea from similar environments, we were expecting our amplification products to be haloarchaeal. Haloarchaeal sequences were indeed identified from the Poland no. 4 sample (Fig. 4a), although this was not quality controlled halite. However, amplification products from the other samples proved to be bacterial sequences, none of which matched known organisms, although these were clearly derived from members of the β - and γ -proteobacteria. This was an unexpected result, considering that the first round of PCR used two archaeal primers and the second round used one archaeal and one universal primer. It has been reported²⁰ that primers need a minimum of three homologous nucleotides at the 3' end for successful priming, and that the rest of the primer needs very little homology for the target. The bacterial sequences obtained with archaeal/universal primers are consistent with this observation. The interplay of primer structure/concen-

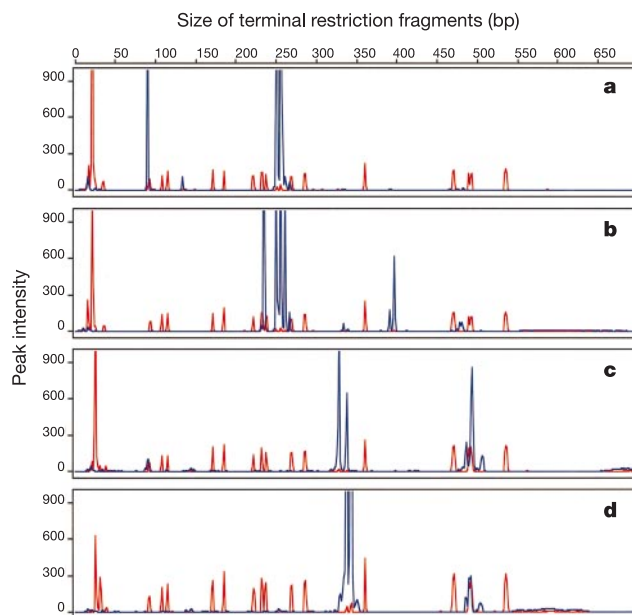


Figure 3 t-RFLP analysis of quality controlled Poland (**a, c**) and Thailand (**b, d**) PCR products generated using FAM-labelled primer FD3 and 1392 Ru. These products were digested with *MspI* (**a, b**) and *RsaI* (**c, d**), and the 5'-terminal fragments resolved using an

ABI 377 automated sequencer. Blue peaks signify PCR products; red indicate calibration peaks. Units of intensity are arbitrary.

tration, annealing temperature and Mg^{2+} concentration could also have influenced the results^{21,22}.

As all samples were treated in the same way, laboratory contamination would have resulted in the same DNA being present in all samples. This was not the case. We did not get PCR product from some samples (Fig. 2), and the PCR products from the others were different on the basis of t-RFLP profiles (Fig. 3a–d). There were, however, three t-RFLP fragments common to Poland no. 4 and Thailand no. 541 (Fig. 3a, b).

One of the closest known relatives to the clones in cluster A (Fig. 4b) is the β -proteobacterium, *Aquabacterium commune* (98.4–99.0%), isolated from a drinking-water biofilm²³. Salt-crystal clones in cluster A and the aforementioned species are closely related to *Leptothrix* species, and so it is worth noting that abundant structures in ancient halite of what appeared to be *Leptothrix*-like iron-oxidizing bacteria have been reported²⁴.

The remaining salt-crystal clones belong to the γ -proteobacteria. Cluster B is represented by clones from Poland no. 4 and clusters C and D are represented by clones from Thailand no. 541 (Fig. 4b). The three clones in cluster D belong to the physiologically diverse *Pseudomonas fluorescens* group. In cluster B, clones Poland no. 16, 6 and 11 were 99.4%, 98.8% and 98.5% similar, respectively, to *Stenotrophomonas maltophilia*, a phylogenetically broad species that contains over 100 characterized strains isolated from many different environments. The two clones in cluster C (Thailand no. 7 and Thailand no. 14) are 99.5% and 98.6% similar to an *Acinetobacter* species, respectively. Other close relatives include environmental clones from the mud in the Marianas Trench, the deepest

part of the ocean (99.4% and 98.6%, respectively), and from high-temperature, moderately saline oil reservoirs (99.0% and 98.2%, respectively).

The close relatives of these salt-crystal phylotypes are thus ubiquitous, and live in a wide variety of environments, including the subsurface. Therefore, it is probable that some of these related but geographically distinct organisms have been separated for millions of years, yet they still share very similar 16S rRNA sequences. This lends support to the argument that the molecular clock may be slow in certain phylogenetic lineages. Accordingly, it would be unwise to ascribe undue significance to gene sequence comparisons between salt-crystal phylotypes and the 16S rRNA genes of current isolates—sequence similarities do not necessarily rule out the possibility that the salt-crystal phylotypes represent the signatures of organisms that were trapped in the salts as they were deposited millions of years ago.

The haloarchaeal sequences from the Polish halite are all closely related but not identical (Fig. 4a). A marked difference is seen between the amplicons obtained from quality controlled and non-quality controlled Polish halite. From the former, only bacterial sequences were obtained; however, from the latter, haloarchaeal, in addition to bacterial sequences were obtained. The difference between the two may relate to the condition of the halite. The halite that was not quality controlled contained visible open and healed microfractures. Of note, the haloarchaeal sequences obtained match those of haloarchaea related to but distinct from the surface isolate *Halobacterium salinarum*, which are known to inhabit brine pools associated with mined Permian–Triassic evaporites⁴.

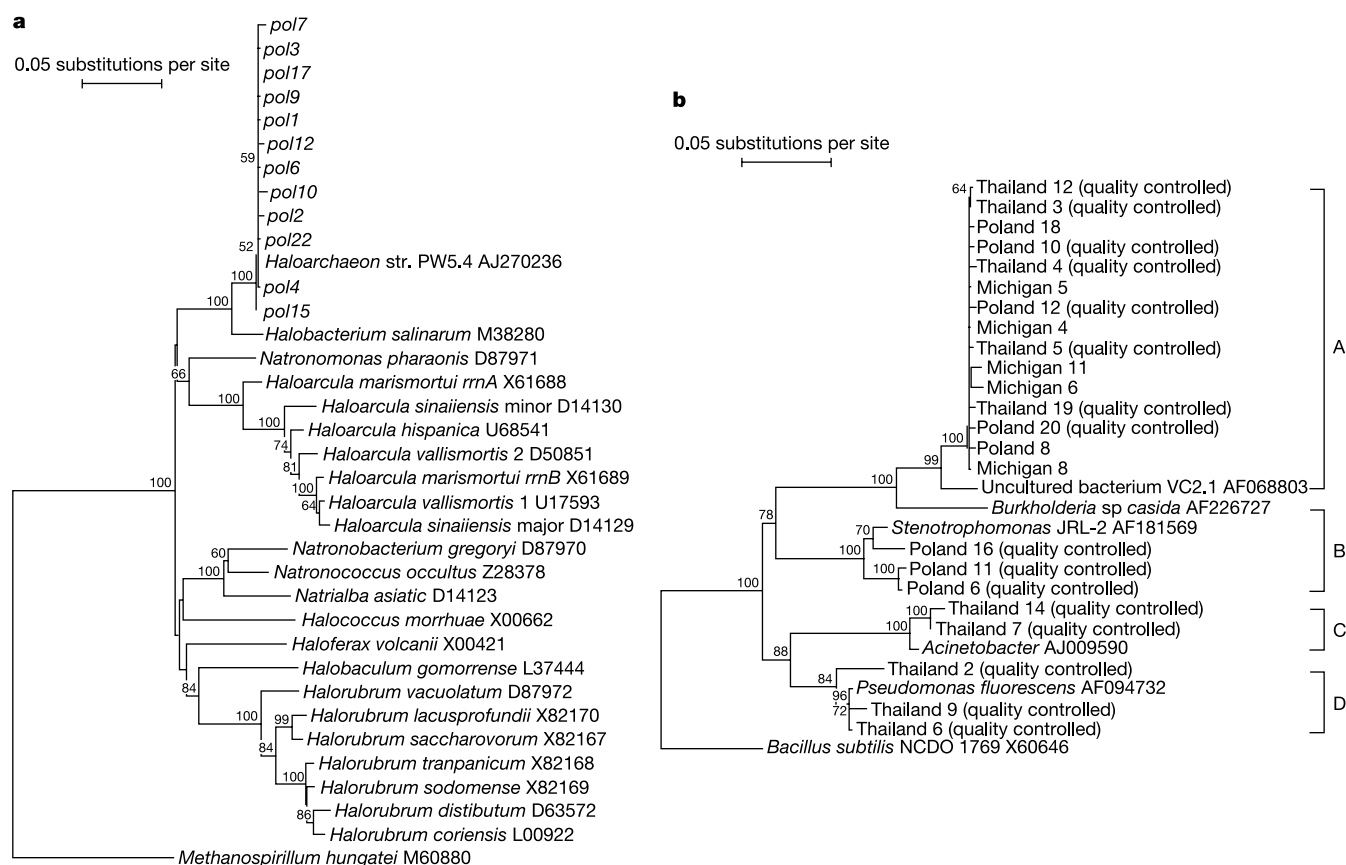


Figure 4 Phylogenetic trees of halobacterial (a) and bacterial (b) 16S rRNA genes obtained from ancient halite samples from Poland, Thailand and Michigan prepared from quality controlled and non-quality controlled material. The bootstrapped trees were drawn using the distance matrix method of ref. 28 with the neighbour-joining method of ref. 29 to

infer tree topology. Bootstrap values of greater than 50% are shown. The four clusters in b (A–D) are indicated on the right. Accession numbers of sequences for known organisms and phylotypes are indicated in the trees.

We failed to recover viable cells in a limited number of growth experiments involving excised quality controlled halite pieces despite optimizing growth medium (data not shown). The main factor mitigating against recovery is probably the small sizes of the pieces, as we have previously shown that relatively large amounts of halite are necessary to guarantee recovery of viable cells in enrichment experiments²⁵. Heat transfer to the centres of small pieces would also be rapid. Our methods of testing primary fluid composition cannot be used to screen anything other than small pieces.

We are uncertain whether the DNA isolated from the primary halite crystals derives from within fluid inclusions or from within the crystalline halite. Nor do we know whether the DNA originated from intact microorganisms or whether it is free DNA from lysed cells. The DNA might also be derived from non-halophilic, moribund bacteria or from halophilic microorganisms living in saturated brine that became entrapped as the halite crystals formed. We conclude that DNA entrapped in halite can survive over geological timescales—we have previously discussed the special nature of the hypersaline environment as a medium for the possible preservation of biological signatures^{8,26}. □

Methods

Sample preparation

Except for sample Poland no. 4 that was obtained from a salt mine in Poland, the samples were obtained exclusively from bore hole cores recovered over many years (2–20 years) by the University of Binghamton, using conventional diamond drilling at depths up to 1 km below the present-day surface. The cores, once recovered, were stored under dry conditions at ambient temperatures. Pieces were cut from the cores during the last 2–3 years and again stored under dry conditions at ambient temperatures.

Thin slices (1–2 mm thickness) were cut from the original bulk salt samples using a micro-saw and then hand ground and polished on 8.4-μm silicon carbide abrasive paper under propan-2-ol to remove the surface layer of salt. We used new paper for each sample. Pieces were then ultrasonically cleaned in propan-2-ol to remove salt dust and other particulates, washed in fresh propan-2-ol and air dried.

Using 1–2-mm thick slices containing very high densities of primary brine inclusions (typically 10^{2–3} inclusions per mm²), 2–5 mm² portions were carefully cut away using a 532 nm laser. Each sample was placed within a disposable, plastic microdish that was isolated from the laser lens by a disposable, thin glass window. This configuration ensured that no two samples were in contact with the same surface and guaranteed the lowest risk of cross contamination from ablated salt. The samples were then stored in sterile plastic tubes.

To assess the possibly damaging effects of laser radiation on DNA, cells of the halobacterium *Halorubrum vacuolatum* were suspended in the salts component of a standard growth medium within capillary tubes laser irradiated at 266 and 532 nm. The effects were determined by the yield of amplified 16S rDNA from these cells, using PCR (data not shown). Predictably, 266 nm had the more detrimental effect and therefore 532 nm was used for all subsequent manipulations.

DNA extraction

Chips of halite were soaked in 100% ethanol for 30 min. The ethanol was burnt off and the halite held in a Bunsen flame for a further 2 s. The material was extracted using a protocol modified from ref. 27, making sure that enough lysis buffer (consisting of 4 M guanidine isothiocyanate, 50 mM Tris-HCl, pH 7.5, 25 mM EDTA and 1.3% Triton X-100) was added to ensure solubilization. We carried out all manipulations in a positive-pressure HEPA air-filtered, clean room.

PCR amplification

Nested PCR amplification of 16S rRNA genes from both sets of samples were carried out in a dedicated PCR cabinet. Nested PCR used primers 27Fa 5'-TC(C/T)GGTTGATCCTG(C/G)CGG-3' and 1524Ra 5'-AGGAGGTGATCCAGCC-3' in the first round reaction, with 353Fa 5'-GGCCCTACGGGGCGCA-3' and 1392Ru 5'-ACGGGCGGTGTGT(A/G)C-3' in the second round reaction. The thermal cycling programme was 94 °C for 4 min 30 s, followed by 30 cycles of 94 °C for 30 s, touchdown for 30 s from 62 °C with a 0.5 °C decrease in temperature per cycle, and 72 °C for 2 min. A final extension step was used for 10 min at 72 °C.

Terminal restriction fragment length polymorphism

The fluorescently labelled PCR products from Poland and Thailand quality controlled samples for the t-RFLP analysis were generated using step 1 above, with a second round reaction using the FAM-labelled forward primer FD3 5'-AAACT(C/T)AAA(G/T)GAATTGACGG-3' with 1392Ru. PCR products were digested with the restriction endonucleases *MspI* and *RsaI*. The resulting fluorescent terminal fragments were resolved and analysed using an ABI 377 automated sequencer and Gene Scan software 2.1.1.

Sequencing

Amplified DNA was cloned into *Escherichia coli* using the vector pGEMT-Easy. Plasmid preparations were made and sequences from both strands obtained using M13 forward and reverse sites in the vector, coupled with internal sites to the cloned DNA. These

sequences were compared to other rDNA sequences using FastA and subsequently converted to a tree format using Genedoc and Treecon software (see Fig. 4a, b). Sequencing primers were: M13 forward and reverse, 806Fu 5'-ATTAGATACCC(T/G)GGTAGTCC-3', FD3 5'-AAACT(C/T)AAA(G/T)GAATTGACGG-3', RD3 5'-CCGTCAATTC(A/C)TTT(G/A)AGTTT-3' and RD2 5'-G(T/A)ATTACGCGGC(G/T)GCTG-3'.

Received 1 June 2001; accepted 7 March 2002.

- Woodward, S. R., Weyand, N. J. & Bunnell, M. DNA sequence from Cretaceous period bone fragments. *Science* **266**, 1229–1232 (1994).
- Cano, R., Poinar, H. N. & Poinar, G. O. Isolation and partial characterisation of DNA from the bee *Proplebeia dominicans* (Apidae: Hymenoptera) in 25–40 million year old Dominican amber. *Med. Sci. Res.* **20**, 249–251 (1992).
- Smith, A. & Austin, J. J. Can geologically ancient DNA be recovered from the fossil record? *Geoscientist* **7**, 8–11 (1998).
- McGenity, T. J., Gemmell, R. T., Grant, W. D. & Stan-Lotter, H. Origins of halophilic microorganisms in ancient salt deposits. *Environ. Microbiol.* **2**, 243–250 (2000).
- Vreeland, R. H., Rosenzweig, W. D. & Powers, D. W. Isolation of a 250 million year old halotolerant bacterium from a primary salt crystal. *Nature* **407**, 897–899 (2000).
- Hazen, R. M. & Roedder, E. How old are bacteria from the Permian age? *Nature* **411**, 155 (2001).
- Norton, C. F. & Grant, W. D. Survival of halobacteria within fluid inclusions in salt crystals. *J. Gen. Microbiol.* **134**, 1365–1373 (1988).
- Gemmell, R. T., McGenity, T. J. & Grant, W. D. Use of molecular techniques to investigate possible long term dormancy of halobacteria in ancient salt deposits. *Ancient Biomol.* **2**, 125–133 (1998).
- Wardlaw, N. C. & Schwerdtner, W. M. Halite-anhydrite seasonal layers in the Middle Devonian Prairie Formation, Saskatchewan, Canada. *Geol. Soc. Am. Bull.* **77**, 331–342 (1966).
- Roedder, E. The fluids in salt. *Am. Mineral.* **69**, 413–439 (1984).
- Shepherd, T. J., Ayora, C., Cendón, D. I., Chenery, S. R. & Moissette, A. Quantitative solute analysis of single fluid inclusions in halite by LA-ICP-MS and cryo-SEM-EDS: complementary microbeam techniques. *J. Eur. Mineral.* **10**, 1097–1108 (1998).
- Gunther, D., Audetat, A., Frischknecht, R. & Heinrich, C. A. Quantitative analysis of major, minor and trace elements in fluid inclusions using laser ablation-inductively coupled plasma mass spectrometry (LA-ICP-MS). *J. Anal. At. Spectrom.* **13**, 263–270 (1998).
- Timofeeff, M. N., Lowenstein, T. K. & Hardie, L. A. in *Ninth Ann. V.M. Goldschmidt Conf.* 299 (abstract) (Lunar and Planetary Institute, Houston, 1999).
- Hardie, L. A. Evaporites, marine or non-marine? *Am. J. Sci.* **284**, 193–240 (1984).
- Hardie, L. A. Secular variations in seawater chemistry: an explanation for the coupled secular variation in the mineralogies of marine limestone and potash evaporites over the past 600 my. *Geology* **24**, 279–283 (1996).
- Zimmermann, H. On the origin of fluids included in Phanerozoic marine halite; basic interpretation strategies. *Geochim. Cosmochim. Acta* **65**, 35–45 (2001).
- Timofeeff, M. N., Lowenstein, T. K. & Blackburn, W. H. ESEM-EDS. An improved technique for major element chemical analysis of fluid inclusions. *Chem. Geol.* **164**, 171–182 (2000).
- Timofeeff, M. N. *et al.* Evaluating seawater chemistry from fluid inclusions in halite: examples from modern marine and non-marine environments. *Geochim. Cosmochim. Acta* **65**, 2293–2300 (2001).
- Rosler, H. J. & Lange, H. *Geochemical Tables—Table 97: Composition of sea water* (Elsevier, Amsterdam/London/New York, 1972).
- Sommers, R. & Tautz, D. Minimal homology requirements for PCR primers. *Nucleic Acids Res.* **17**, 674 (1989).
- Rychlik, W. Priming efficiency in PCR. *Biotechniques* **18**, 84–89 (1995).
- Rychlik, W., Spencer, W. J. & Rhoads, R. E. Optimization of the annealing temperature for DNA amplification *in vitro*. *Nucleic Acids Res.* **18**, 6409–6412 (1990).
- Kalmbach, S., Manz, W., Wecke, J. & Szewzyk, U. *Aquabacterium* gen. nov., with description of *Aquabacterium citratiphilum* sp. nov., *Aquabacterium parvum* sp. nov., and *Aquabacterium commune* sp. nov., three *in situ* dominant bacterial species from the Berlin drinking water system. *Int. J. Syst. Bacteriol.* **49**, 769–777 (1999).
- Müller, A. & Schwartz, W. Über das Vorkommen von Mikroorganismen in Salzlagernstätten (geomikrobiologische Untersuchungen III). *Zeitschrift Deutschen Geologischen Gesellschaft* **105**, 789–802 (1953).
- Norton, C. F., McGenity, T. J. & Grant, W. D. Archaeal halophiles (halobacteria) from two British salt mines. *J. Gen. Microbiol.* **139**, 1077–1081 (1993).
- Grant, W. D., Gemmell, R. T. & McGenity, T. J. Halobacteria: the evidence for longevity. *Extremophiles* **2**, 279–287 (1998).
- Höss, M. & Pääbo, S. DNA extraction from Pleistocene bones by a silica-based purification method. *Nucleic Acids Res.* **21**, 3913–3914 (1993).
- Jukes, T. H. & Cantor, C. R. in *Mammalian Protein Metabolism* (ed. Munro, H. N.) Vol. 3, 21–132 (Academic, New York, 1969).
- Saitou, N. & Nei, M. The neighbour-joining method: a new approach for constructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406–425 (1987).

Acknowledgements

The samples were provided by M. N. Timofeeff, S. Brennan and T. K. Lowenstein, and we acknowledge their contribution to the research. We also thank D. Gunther and B. Hattendorf for undertaking corroborative chemical analysis of the brine inclusions. This work was supported by an NERC grant.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to W.D.G. (e-mail: wdg1@le.ac.uk) or T.J.S. (e-mail: t.shepherd@bgs.ac.uk). Sequences described in this paper are deposited in GenBank under accession numbers AJ319543–AJ319577.

Coexistence and relative abundance in forest trees

Colleen K. Kelly*†‡ & Michael G. Bowler‡§

* Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK

§ Department of Physics, University of Oxford, Keble Road, Oxford OX1 3RH, UK

‡ These authors contributed equally to this work

Contemporary acceleration of biodiversity loss makes increasingly urgent the need to understand the controls of species coexistence^{1,2}. Tree diversity in particular plays a pivotal role in determining terrestrial biodiversity, through maintaining diversity of its dependent species^{3,4} and with them, their predators and parasites. Most theories of coexistence based on the principle of limiting similarity suggest that coexistence of competing species is inherently unstable; coexistence of competitors must be maintained by external forces such as disturbance^{5,6}, immigration⁷ or 'patchiness' of resources in space and time^{8,9}. In contrast, storage theory postulates stable coexistence of competing species through temporal alternation of conditions favouring recruitment of one species over the other^{10,11}. Here we use storage theory to develop explicit predictions for relative differences between competitors that allow us to discriminate between coexistence models. Data on tree species from a primary forest on the Mexican Pacific coast support a general dynamic of storage processes determining coexistence of similar tree species in this community, and allow us to reject all other theories of coexistence.

A central dictum of community ecology stipulates that coexistence of similar species is transient, with the more abundant competitive superior eventually excluding its inferior^{12,13}. Nonetheless, highly similar species coexist. A recent comparison of pairs of closely related, ecologically similar tree species showed the more abundant to have 'smoother' population size profiles than the less abundant species at the same site¹⁴. Such deviations from a smooth curve signify recruitment fluctuation, suggesting storage processes. (Storage here refers to the ability to recover from periods of low recruitment being 'stored' in the long-term survival of reproductive capacity in long-lived adults or a seed bank¹⁵.) Storage theory predicts that an individual species will show greater fractional deviation in recruitment at low than at high abundance^{10,11}, but has not explicitly addressed the relative relationships to be expected of both sides of a resident pair of competitive tree species. Here we develop a model to expose the role of storage dynamics in the coexistence of ecologically similar, closely related tree species in nature. Using information from a set of coexisting tree species native to the tropical deciduous forest of Mexico, we find a generalizable pattern supporting the predictions made by our model of greater fractional deviation and competitive advantage in the less abundant of two similar taxa. Either one or both of these predictions contrasts with the expectations of all other current models of species coexistence^{6,7,9,13}.

We suppose that when a site that can support an adult tree becomes available, there is some probability that it is most suitable for species in class A (A might be a genus), and that such sites are colonized at once by species a_i which are members of class A. We have in mind spatio-temporal partitioning into such site classes^{8,9}. In a mature forest let a fraction p_1 of the A sites be occupied by a_1 ,

and p_2 by a_2 . We write

$$\frac{dp_i}{dt} = -\delta_i p_i + R_i(t) \quad (1)$$

where δ_i is the death rate parameter and R_i is the probability of recruitment in unit time. Equation (1) integrates to the asymptotic solution

$$p_i(t) = \int_0^\infty R_i(t-g)e^{-g\delta_i} dg \quad (2)$$

where g is the time elapsed since recruitment. The function

$$\Pi_i(t, g) = R_i(t-g)e^{-g\delta_i} \quad (3)$$

is the age profile of species a_i . The present discussion is limited to the case where $\delta_1 = \delta_2 = \delta$. This assumption simplifies the treatment, and is consistent with such data as exist for the species we consider.

If class A vacancies are filled at once by species belonging to A, then in a mature forest in which species a_1 and a_2 coexist in the long term:

$$R_1 + R_2 = \bar{R}_1 + \bar{R}_2 \quad (4a)$$

$$\bar{p}_i = \frac{\bar{R}_i}{\delta} \quad [\bar{p}_1 + \bar{p}_2 = 1] \quad (4b)$$

Overbar indicates long-term average. If the recruitment rate R_1 fluctuates, then R_2 responds so as to maintain $R_1 + R_2 = \delta$. As an example, R_1 might go to zero on a quasiperiodic basis as a result of environmental fluctuations. Regardless of detail, because $\Delta R_1 = -\Delta R_2$, the fractional fluctuations in recruitment $\Delta R/\bar{R}$ are larger for the species with the smaller mean recruitment, and from equation (2) the species with the smaller mean recruitment is the rarer, given similar death rates. The greater the disparity in abundance, the greater the ratio of fractional fluctuations (because the absolute fluctuations in recruitment are of equal magnitude).

One might expect the closely related species a_1 and a_2 , competing for the same sites, to relax to a configuration in which one is wholly dominant^{12,13}, but recruitment fluctuations can permit coexistence. A model that satisfies the assumptions we have made is the two-component lottery model¹⁰, in which if juveniles of a_1 are quasiperiodically destroyed, a_1 and a_2 can coexist provided that a_1 has sufficient competitive advantage during the periods in which the young a_1 are safe. The necessary assumptions in this treatment do not have to be extreme; the relationship between competitors will exist if A type sites are colonized by class A faster than by species in another class and if a_2 can colonize vacant A sites in a time short in comparison with the typical length of an a_1 crisis.

In this two-component lottery model, a location suitable for a new adult becomes available at a rate equal to the death rate of adults, and the probability of species a_1 taking it is

$$\frac{\beta_1^* p_1}{\beta_1^* p_1 + \beta_2^* p_2} \quad (5)$$

where β^* is the product of a birth rate multiplied by a competitive factor (see below), p_1 is the fraction of sites occupied by a_1 , and $p_1 + p_2 = 1$ at all times in a mature system.

Then

$$\frac{dp_1}{dt} = -\delta p_1 + \frac{\delta p_1 \beta_1^*}{\beta_1^* p_1 + \beta_2^* (1 - p_1)} \quad (6)$$

where the second term on the right is recruitment R_1 .

On very general grounds we look to fluctuations in recruitment for the principal cause of fluctuations in the age profiles Π_i , and in a mature forest fluctuations in recruitment correspond to fluctuations in the β^* parameters. Suppose that the parameter β_1^* is quasiperiodically zero for a spell (because seedlings get washed out or

† Present address: Division of Biodiversity and Ecology, School of Biological Sciences, University of Southampton, Bassett Crescent East, Southampton SO16 7PX, UK

eaten or baked), but during an environmentally favourable period it is constant at β_2^* . Suppose that the other species is less sensitive to these environmental effects and maintains (at least approximately) a constant value β_1^* . (We note that β_2^* can be suppressed without affecting the results, provided that $\beta_1^*/\beta_2^* \ll 1$ during unfavourable periods.) During periods with β_1^* 'off', $R_1 = 0$ and $R_2 = \delta$, regardless of the value of β_2^* . If β_1^* is 'on' for a fraction f of the time and both populations are in stable long-term coexistence:

$$\bar{p}_1 = \frac{f - (\beta_2^*/\beta_1^*)}{1 - (\beta_2^*/\beta_1^*)} \quad (7)$$

It follows that $\beta_1^*/\beta_2^* < f$ species a_1 must have a competitive advantage during the favourable periods. This agrees with the separately derived result of ref. 16. If $\beta_1^* \gg \beta_2^*$ then $\bar{p}_1 \rightarrow f$; if $f = 1/2$, $\bar{p}_1$ only reaches 1/2 (equal abundance) if species a_1 has an overwhelming competitive advantage during favourable periods. For $f \leq 1/2$, species a_1 (which suffers fluctuations in β_1^*) is inevitably the rarer, yet has a competitive advantage. It is of course possible for a_1 to be the commoner; the condition for $\bar{p}_1 > 1/2$ is

$$\frac{\beta_2^*}{\beta_1^*} < 2f - 1 \quad (8)$$

There is thus a wide range of circumstances in which the rarer species is the one that suffers recruitment crises and which must possess a significant competitive advantage during favourable periods in order to survive at all (Fig. 1). A growth rate/herbivore defence trade-off is a biologically plausible example of this¹⁷. In the absence of recruitment crises, species a_1 would drive out a_2 . The treatment given here specifically addresses long-lived species, but similar relationships may exist for shorter-lived organisms possessing a holding stage or 'bank' for non-adults (compare ref. 11).

Although this lottery model predicts that fluctuations will be in strict antiphase, under natural conditions that need not be the case and storage processes will still function. If the number of sites available varies with time, then an in-phase variation in recruitment is superimposed on the antiphase behaviour. Both species suffer a component of fluctuation in recruitment (by number) that is the same fraction of mean recruitment for each.

We predict from the above model that (1) the rarer species will show greater fractional deviation in recruitment fluctuations than the more common, and (2) the rarer species will show a competitive advantage over the more common. We have tested these hypotheses using a subset of species from a larger field project investigating relative abundance¹⁴. The project design of multiple congeneric

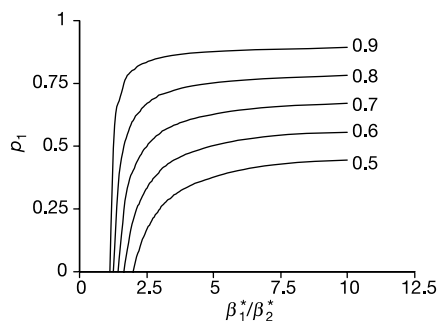


Figure 1 Curves for values of f , the fraction of time in which recruitment for species 1 is 'on'. β_1^*/β_2^* is the ratio of the competitive advantage of species 1 over species 2 when species 1 is 'on'; p_1 is the proportion of species 1 relative to the total number of individuals for both species. The figure illustrates equation (8) in the text, showing that unless the competitive advantage exceeds something like a factor of 2, it is almost impossible for any fluctuations, detectable in the population profile data to allow species 1 (which is 'off' for $1 - f$) to be commoner than species 2.

species pairs matched for ecological similarity as well as for the shared genetic background of close phylogenetic relationship maximizes the degree of competition between paired taxa^{12,13} while simultaneously minimizing potentially confounding variables extraneous to the active processes of coexistence^{18,19}. Nine of the 12 species of the larger project had the growth rate data necessary to test our hypotheses. These data supported five congeneric contrasts for each prediction²⁰, allowing a statistical test of whether storage dynamics may be inferred to be a general process in this forest²¹.

We compared recruitment fluctuations in the target species by calculating the deviation of the age profile of canopy level individuals of each species from an exponential curve parameterized by the mean age value for that species. An exponential distribution offers a smoothly descending shape that would obtain in the absence of recruitment fluctuations, given a death rate independent of the age of canopy level trees²². We found that in all five contrasts, the rarer taxa of the pair showed a greater fractional deviation of its age

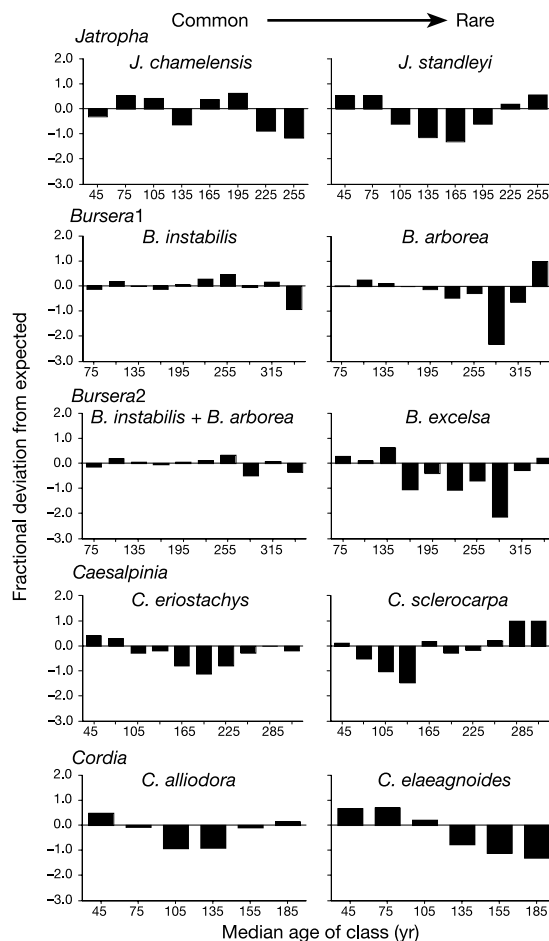


Figure 2 Fractional deviation of age classes from an expected exponential distribution. Note that the deviations shown here were calculated as (expected — observed)/expected, and so differ from the calculation used to estimate the fit of observed to expected distributions (see Methods). The calculation used in this figure better portrays the visual impact of fractional deviation of the individual populations. Comparisons (contrasts) were made within, not between, 'rows' of taxa, which represent pairs of taxa more closely related to one another than to any other taxon in the figure. Field growth measures of *B. instabilis* indicate that individuals of less than 3 cm diameter at breast height (d.b.h.), the assumed minimum d.b.h. for canopy individuals in this study, could be greater than 30 years old, necessitating the use of the 60–90 yr class as the smallest age class for the two contrasts using this genus.

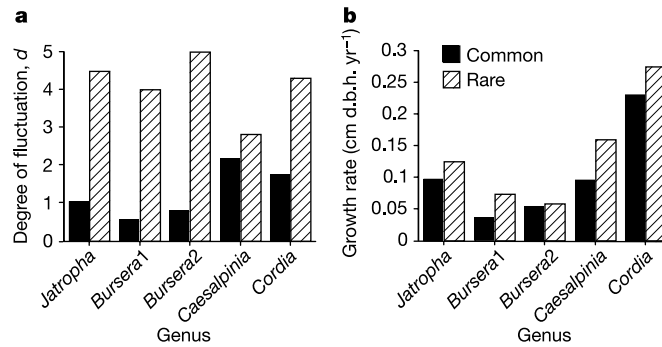


Figure 3 Contrasts of common and rare congeneric taxa. **a**, Growth rates. **b**, Degree of fractional deviation (d). For all five pairs, the relationship between the traits and relative abundance is that predicted by the model presented here. The value of P for each set of

contrasts is 0.031. The probability that the overall results for both sets of contrasts could be due to chance is 0.00096. d.b.h.

profile from the corresponding smooth curve ($P = 0.031$; Figs 2 and 3a). We verified that differences in sampling intensity (the rarer species is represented by a greater proportion of its population) did not affect our conclusions. This was done by comparison with a null model in which we randomly lowered sampling intensity for the more sampled species. In no instance did differences in sampling intensity produce a significantly different fractional deviation value ($P \geq 0.80$ in all cases), or change the relationship within a contrast.

For plants, competitive advantage is revealed mechanistically as greater individual growth rate, signifying either greater resource uptake efficiency or as an advantage in itself in competition for light¹³. Differences in growth rate between congeneric taxa paralleled those found for fractional deviations, with all five contrasts showing less abundant species having higher growth rates than more abundant species (testing against a hypothesis of randomness, $P = 0.031$; Fig. 3b). The two hypotheses of fractional deviations and competitive advantage are independent of one another, and may therefore be combined; the probability of obtaining the observed unanimous support for our predictions in 0.00096.

The underlying principle of storage theory is that of covariance between the environment and competition²³. Many models embodying this principle with environmental fluctuations predict that the rarer of two competing species will exhibit the larger fractional recruitment fluctuations^{11,15}; ours is only one such model. The prediction that the lower density species has the higher mean β^* is not, however, a general prediction of storage theory, but rather derives from the assumption of the greater sensitivity of one of the two taxa to the operative environmental variable.

Of existing models postulating coexistence of competitors, only storage theory explains our results. Models predicating coexistence through resource 'patchiness' and escape from competition^{8,9} offer no expectations of any regular difference in the community in age profiles for more and less abundant competitors, nor does a model assuming coexistence through chance immigration⁷. On the other hand, disturbance theory addresses population fluctuations in competing species by raising the expectation of in-phase pulses, but offers no predictions of relative differences in the degree of those fluctuations^{5,6}. Simultaneity of fluctuations is not observed for any of the paired comparisons ($P \geq 0.7$). On the contrary, visual inspection of the pattern of deviations (Fig. 2) suggests antiphase fluctuations, an expectation of storage theory in the absence of temporal variation in the number of sites. However, this impression was not supported statistically ($0.055 \leq P \leq 0.623$).

Models based on either habitat heterogeneity or chance similarly expect no regular pattern of differences in growth rates such as we found. Although a homogeneous habitat may be associated with a regular pattern of growth advantage, it is one in which the

competitive superior is also the more abundant^{5,6,13,24}. Thus we may doubly reject these alternative models of coexistence.

We conclude by emphasizing that our comparative pairs demonstrate storage dynamics without exception, even though the target taxa were not selected with any *a priori* expectation of demonstrating this process. With our test cases drawn from widely separated plant orders and subclasses, our results therefore offer strong evidence for storage dynamics as a general process determining the coexistence of tree species²⁵. This conclusion has two immediately important practical applications. First, maintenance of tree diversity requires, in addition to the expected identification and maintenance of conditions that allow physical establishment and reproduction, identification and maintenance of the factors causing particular sorts of temporal fluctuation in those processes. This conclusion is made urgent for general biodiversity management through the role that tree diversity plays in supporting animal diversity^{3,4}. Second, our findings of recruitment cycles of decades to centuries show population matrix analyses to hold little utility for tree species; the results of such analyses, even those based on a decade or more of information^{22,26}, will depend on the segment of the fluctuation cycle captured rather than the stability, or otherwise, of the population. □

Methods

Chamela Biological Station comprises 1,600 ha, and is centred at 19° 13' N and 105° 03' W²⁷. Vegetation is primarily tropical deciduous forest, with bands of tropical semi-deciduous forest in the semi-seasonal drainage streams²⁸. We use the term 'abundance' with regard to local abundance, and compared between congeners as average numbers of stems per hectare (from ref. 14).

Growth rates were calculated from change in diameter of individuals assayed in 1987 and again in 1998 or 1999, all occurring in seven permanent 30 m × 80 m plots established in 1987. For multi-stemmed individuals, cross-sectional area for each stem was calculated from diameter at breast height (d.b.h.) summed and back-calculated to give a single size value for the individual.

Samples of 120 to 500 individuals were located and measured over the course of the rainy seasons of 1998 and 1999. Exhaustive censuses were conducted within defined areas outside the permanent plots, conferring a similarity in distribution between conspecific individuals that permitted conversion of size into age through dividing d.b.h. by growth rate. Size is a reasonable correlate of age in tropical trees²⁹ if growth is not greatly dependent on initial size, and we found no significant relationships between initial size and growth rate ($P \geq 0.27$). Population profiles were compared between taxa as age distributions of individuals ≥ 30 years of age, except for *Bursera* for which 60 years was the minimum age. Minimum age for inclusion into the age profile was that which could be assumed to have reached the canopy of this low forest (4–12 m) and allowed a minimum of six classes, thus permitting a sign test of non-directional hypotheses²¹.

Use of taxon-specific curves to determine degree of fluctuation for a taxon minimized error arising from misestimation of ages of individuals¹⁴. Deviations of observed from expected distributions (d) were calculated as

$$d = \sum \left[\frac{(N_i^{\text{expected}} - N_i^{\text{observed}})^2}{N_i^{\text{expected}^2}} \right] - \sum 1/N_i^{\text{expected}} \quad (9)$$

with N_i^{observed} representing the number of individuals in the i th category of the observed distribution, and N_i^{expected} representing the number of individuals expected in the i th category of the expected distribution (see Fig. 2). This expression best estimates

$\sum (\Delta R/\bar{R})^2$, which the model predicts will be greater for less abundant species, and allows for differences in lifetimes and sample sizes. The second term corrects for merely statistical fluctuation, assuming Poisson statistics. For each contrast, the estimator was summed over age classes for each species.

Bursera instabilis is more closely related to *Bursera arborea* than either is to *Bursera excelsa*³⁰. Therefore a higher level contrast was formed to compare *B. excelsa* to a single value of the two other congeners (*Bursera* 2). A weighted pooling of the data for *B. instabilis* and *B. arborea* was based on the abundance of the former relative to the latter, drawing on the ecological assumption that the more distantly related *B. excelsa* would interact with the two other, closely related species as a single functional entity.

Samples represented a larger proportion of the populations of the rarer than the more common species. We therefore compared our observed results with a null model in which a subsample of a size corresponding to the sampling intensity of the less heavily sampled species was randomly selected without replacement from the list of age values for the more heavily sampled species. Deviation of this subsample from a smooth curve was determined as for the observed populations, repeated 10⁴ times. In all five instances, the null model showed that for the less abundant, more intensively sampled species, a less intensive sampling regime produced a value that did not differ significantly from observed (analysis of variance; $P \geq 0.80$ in all cases).

We compared synchrony within contrast pairs in fluctuation patterns in a sign test, asking whether age classes of comparable taxa deviated in concert from the expected distribution for the taxon significantly more often than expected by chance.

Received 18 September 2001; accepted 11 February 2002.

- May, R. M., Lawton, J. H. & Stork, N. E. in *Extinction Rates* (eds Lawton, J. H. & May, R. M.) 1–24 (Oxford Univ. Press, Oxford, 1995).
- Tilman, D. Causes, consequences and ethics of biodiversity. *Nature* **405**, 208–211 (2000).
- May, R. M. How many species are there on Earth? *Science* **241**, 1441–1449 (1988).
- Kelly, C. K. & Southwood, T. R. E. Species richness and resource availability: a phylogenetic analysis of insects associated with trees. *Proc. Natl Acad. Sci. USA* **96**, 8013–8016 (1999).
- Hutchinson, G. E. Homage to Santa Rosalia, or Why are there so many kinds of animals? *Am. Nat.* **93**, 145–159 (1959).
- Connell, J. H. Diversity in tropical rainforests and coral reefs. *Science* **199**, 1302–1310 (1978).
- Hubbell, S. P. *The Unified Neutral Theory of Biodiversity and Biogeography* (Princeton Univ. Press, Princeton, 2001).
- Nee, S. & May, R. M. Dynamics of metapopulations: habitat destruction and competitive coexistence. *J. Anim. Ecol.* **61**, 37–40 (1992).
- Palala, S. W. & Tilman, D. Limiting similarity in mechanistic and spatial models of plant competition in heterogeneous environments. *Am. Nat.* **143**, 222–257 (1994).
- Chesson, P. L. & Warner, R. R. Environmental variability promotes coexistence in lottery competitive systems. *Am. Nat.* **117**, 923–943 (1981).
- Chesson, P. L. & Huntly, N. Short-term instabilities and long-term community dynamics. *Trends Ecol. Evol.* **4**, 293–298 (1989).
- MacArthur, R. H. & Levins, R. The limiting similarity, convergence and divergence of coexisting species. *Am. Nat.* **101**, 377–385 (1967).
- Tilman, D. *Plant Strategies and the Dynamics and Structure of Plant Communities* (Princeton Univ. Press, Princeton, 1988).
- Kelly, C. K. *et al.* Investigations in commonness and rarity: a comparison of size distributions of co-occurring, congeneric Mexican trees. *Ecol. Lett.* **4**, 618–627 (2001).
- Chesson, P. L. Mechanisms of maintenance of species diversity. *Annu. Rev. Ecol. Syst.* **31**, 343–366 (2000).
- Warner, R. R. & Chesson, P. L. Coexistence mediated by recruitment fluctuations: a field guide to the storage effect. *Am. Nat.* **125**, 769–787 (1985).
- Hermis, D. A. & Mattson, W. J. The dilemma of plants: to grow or defend. *Q. Rev. Biol.* **67**, 283–335 (1992).
- Harvey, P. H. & Pagel, M. D. *The Comparative Method in Evolutionary Biology* (Oxford Univ. Press, Oxford, 1991).
- Harvey, P. H. Phylogenies for ecologists: the 1995 Tansley Lecture. *J. Anim. Ecol.* **65**, 255–263 (1996).
- Pagel, M. D. A method for the analysis of comparative data. *J. Theor. Biol.* **156**, 431–442 (1992).
- Siegel, S. *Nonparametric Statistics* (McGraw-Hill, New York, 1956).
- Condit, R., Sukumar, R., Hubbell, S. P. & Foster, R. B. Predicting population trends from size distributions: a direct test in a tropical community. *Am. Nat.* **152**, 495–509 (1998).
- Chesson, P. Multispecies competition in variable environments. *Theor. Popul. Biol.* **45**, 227–276 (1994).
- MacArthur, R. H. On the relative abundance of bird species. *Proc. Natl Acad. Sci. USA* **43**, 293–295 (1957).
- Woodward, F. I. & Kelly, C. K. in *Plant Functional Types* (eds Smith, T. M., Shugart, H. H. & Woodward, F. I.) (Cambridge Univ. Press, Cambridge, 1997).
- Rees, M., Condit, R., Crawley, M., Palala, S. & Tilman, D. Long-term studies of vegetation dynamics. *Science* **293**, 650–655 (2001).
- Lott, E. J. Annotated checklist of the vascular plants of the Chamela Bay region, Jalisco, México. *Occas. Pap. Calif. Acad. Sci.* **148**, 1–60 (1993).
- Rzedowski, J. *Vegetación de México* (Limusa, México DF, México, 1978).
- Lusk, C. H. & Smith, B. Life history differences and tree species coexistence in an old-growth New Zealand rain forest. *Ecology* **79**, 795–806 (1998).
- Becerra, J. X. & Venable, D. L. Nuclear ribosomal DNA phylogeny and its implications for evolutionary trends in Mexican *Bursera* (Burseraceae). *Am. J. Bot.* **86**, 1047–1057 (1999).

Acknowledgements

We thank F. I. Woodward, P. Doncaster, P. Harvey, M. Fenner and P. Chesson for comments. This work was supported by grants to C.K.K. from the US-Mexico Foundation for Science, the Earthwatch Foundation and the National Geographic Society. C.K.K.

thanks Merton College for a visiting research fellowship during the initiation of this project.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to C.K.K. (e-mail: c.k.kelly@soton.ac.uk).

Host-plant adaptation drives the parallel evolution of reproductive isolation

Patrik Nosil*, Bernard J. Crespi* & Cristina P. Sandoval†

* Department of Biosciences, Behavioural Ecology Research Group, Simon Fraser University, Burnaby BC, Canada, V5A 1S6

† Marine Science Institute, University of California, Santa Barbara, California 93206, USA

Parallel evolution of similar traits in independent populations that inhabit ecologically similar environments strongly implicates natural selection as the cause of evolution¹. Parallel speciation is a special form of parallel evolution where traits that determine reproductive isolation evolve repeatedly, in closely related populations, as by-products of adaptation to ecological conditions^{1,2}. The outcome of such parallel evolution is that ecologically divergent pairs of populations exhibit greater levels of reproductive isolation than ecologically similar pairs of populations of a similar or younger age^{2–4}. The parallel evolution of reproductive isolation provides strong evidence for natural selection in the process of speciation¹, but only one conclusive example from nature is known². Populations of the walking-stick insect *Timema cristinae* that use different host-plant species have diverged in body size and shape, host preference, behaviour and the relative frequency of two highly cryptic colour-pattern morphs^{5,6}. Here we report that divergent selection for host adaptation, and not genetic drift, has promoted the parallel evolution of sexual isolation in this species. Our findings represent a clear demonstration that host-plant adaptation can play a crucial and repeatable role in the early stages of speciation.

Timema walking-stick insects are wingless, phytophagous insects that inhabit the chaparral of California, other areas of the western United States, and northern Mexico. Species are monomorphic or polymorphic for colour, presence of a dorsal stripe, or both, and these patterns match those of their host plants^{5–7}. Phylogenetic studies suggest that speciation in this genus has involved specialization on different host plants⁷. *Timema cristinae* exhibits two genetically determined⁸ colour-pattern morphs, with an unstriped morph more common on *Ceanothus spinosus* and a striped morph more common on *Adenostoma fasciculatum* (mean frequency of unstriped morph on *Ceanothus* is 81%, range 33–100%; mean frequency of unstriped morph on *Adenostoma* is 28%, range 0–100%). Predation on *T. cristinae* by birds and lizards is strong and each morph is most cryptic on the host plant on which it is more common^{5,6}. Thus populations of *T. cristinae* using different host-plant species are exposed to intense divergent selection for crypsis. Patches of these two host plant species are distributed in parapatric mosaics, and local morph frequencies are determined by a gene flow–selection balance between host-plant patches exhibiting the different selective regimes⁶.

Local adaptation to host-plant species in *T. cristinae* is also

indicated by a number of other phenotypic traits that are more divergent between pairs of populations using different host plants than between pairs using the same host. These traits include body size (first principal component, PC1: males, Mantel test $t = -1.86$, $P < 0.05$; females, Mantel test $t = -1.49$, $P < 0.05$), body shape (CVA1: males, Mantel test $t = -1.51$, $P < 0.05$; CVA2: females, Mantel test $t = -1.69$, $P < 0.05$; first and second canonical variate axes from population-level discriminant analyses), host preference⁸ and cryptic resting behaviour (per cent resting where visible from above, the side and below, respectively; from *Ceanothus*: 36%, 33%, 31%; from *Adenostoma*: 16%, 70%, 15%; χ^2 , 2 degrees of freedom (d.f.) = 62.34, $P < 0.001$, $n = 537$). Such host-specific differentiation suggests that host-plant adaptation involves divergence in a suite of complex morphological and behavioural traits. However, physiological trade-offs in host-plant use have not been detected, with females of both morphs exhibiting higher fecundity on *Ceanothus*⁸.

Phylogenetic analyses indicate that different populations using the same host plant do not form monophyletic groups (Shimodaira–Hasegawa tests⁹: on *Adenostoma* the difference in ln likelihood = 77.37 for mitochondrial DNA (mtDNA) (cytochrome oxidase I gene (COI)), 22.57 for nuclear DNA (internal transcribed spacer 2 (ITS-2)), both $P < 0.001$; on *Ceanothus* the difference in ln likelihood = 75.49, $P < 0.001$ for mtDNA, 6.45, $P = 0.08$ for nuclear DNA; Templeton tests¹⁰: $P < 0.05$ for all four analyses). Thus the use of different host-plant species in *T. cristinae* has evolved many times, allowing a replicated test of whether sexual isolation has evolved in parallel with ecological divergence in host-plant use.

No-choice mating trials were conducted between all possible combinations of eight allopatric (geographically separated) populations, yielding a total of 28 pairwise comparisons of sexual isolation. The results revealed strong evidence of host-associated sexual isolation: across all populations, walking-stick insects were more likely to copulate if paired with an opposite-sex member from the same host-plant species than if paired with an opposite-sex member from the different host-plant species (Fig. 1; male host by female host interaction, likelihood ratio test (LR) = 23.98, 1 d.f., $P < 0.001$, logistic regression, $n = 1,024$ mating trials). Consistent with the main prediction of the parallel speciation hypothesis, the magnitude of sexual isolation detected between pairs of populations

using different host plants (I_{PSI} , an overall sexual isolation index = 0.24, s.d. = 0.14, $n = 15$) was significantly greater than the magnitude of isolation detected between pairs of populations using the same host plant ($I_{\text{PSI}} = 0.08$, s.d. = 0.10, $n = 13$) (Mantel test $t = 2.24$, $P < 0.05$; Table 1).

The degree of sexual isolation observed depended strongly on which pairs of populations were compared (LR = 13.61, 4 d.f., $P < 0.01$, logistic regression). We tested three hypotheses concerning the causes of this variation. First, interpopulation differences in colour-pattern morph frequency (per cent difference in frequency of unstriped morph) were positively correlated with the magnitudes of sexual isolation observed between pairs of populations ($r = 0.38$, $P < 0.05$, Mantel test). This result strongly suggests that the net strength of divergent selection for host adaptation to which populations are exposed directly influences the actual magnitude of sexual isolation that evolves. Stronger divergent selection will translate into local adaptation and the evolution of higher levels of sexual isolation. Conversely, gene flow between patches will tend to erode local adaptation⁶ and progress towards speciation.

We found that the geographic distance between populations was positively correlated with the degree of sexual isolation detected ($r = 0.37$, $P < 0.05$, Mantel test), which suggests that geographic isolation facilitates the evolution of sexual isolation. We also found that the degree of sexual isolation observed between populations was not correlated with genetic distances between pairs of populations (mtDNA: $r = 0.13$, $P = 0.29$; nuclear DNA: $r = 0.29$, $P = 0.10$, Mantel tests), and that pairs of populations using different host plants were not more genetically divergent from one another (mean genetic distance, mtDNA: 2.80%, s.d. = 0.81; nuclear DNA: 0.88%, s.d. = 0.83) than pairs of populations using the same host plant (mean genetic distance, mtDNA: 3.07%, s.d. = 0.79; nuclear DNA: 1.10%, s.d. = 0.83; Mantel test $t = 0.47$ and 0.76, respectively, both $P > 0.15$). Given that a molecular clock could not be rejected for the best maximum likelihood tree inferred from the mtDNA data set, variability in the length of time available for genetic drift to differentiate same-host versus different-host populations of *T. cristinae* is therefore unlikely to account for the increased sexual isolation detected between ecologically divergent populations.

Our work indicates that divergent natural selection and sub-

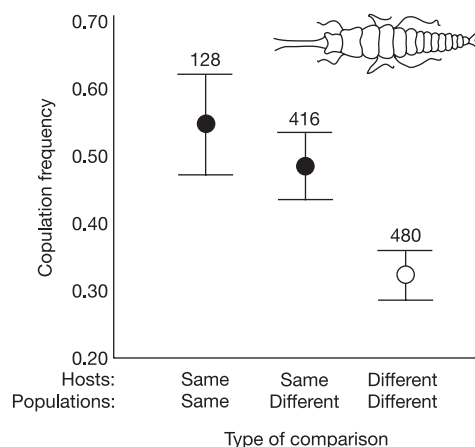


Figure 1 Copulation frequencies were higher for ecologically similar pairs of *T. cristinae* walking-stick insects than for ecologically divergent pairs, where divergence refers to differences in host-plant use ($P < 0.001$, logistic regression). Numbers of mating trials for each pair type are shown above each bar. In between-population mating trials, morph pair-type did not influence the probability of copulation (copulation frequencies when males and females were the same morph and different morphs: different-host crosses,

35%, $n = 219$ and 34%, $n = 261$, respectively, $-2[\log]$ likelihood = 0.18, 1 d.f., $P = 0.67$; same-host crosses, 44%, $n = 218$ and 51%, $n = 198$, respectively, $-2[\log]$ likelihood = 2.01, 1 d.f., $P = 0.17$). This result suggests that colour pattern is not used during mate choice in between-population tests and that host-plant adaptation, rather than selection for crypsis, has directly promoted the evolution of sexual isolation between populations.

Table 1 Levels of sexual isolation between populations of *T. cristinae*

Host comparisons	I_{PSI} (s.d.)	Likelihood ratio test
Different host		
VPC × HV	0.52 (0.19)*	9.99†
VPC × OUTA	0.01 (0.21)	0.01
VPC × L	0.20 (0.21)	2.27
VPC × H	0.11 (0.24)	0.17
VPC × VPA	0.20 (0.21)	1.60
PR × VPA	0.12 (0.20)	1.01
PR × L	0.39 (0.18)*	7.85†
PR × H	0.11 (0.21)	0.31
PR × OUTA	0.36 (0.20)	4.43*
PR × HV	0.20 (0.97)	3.23
P × H	0.19 (0.20)	2.27
P × OUTA	0.16 (0.19)	1.61
P × L	0.39 (0.17)*	9.22†
P × HV	0.33 (0.16)*	7.85†
P × VPA	0.41 (0.18)*	8.09†
Same host		
PR × P	0.08 (0.8)	0.60
PR × VPC	0.00 (0.20)	0.06
P × VPC	0.01 (0.20)	0.07
VPA × OUTA	0.07 (0.19)	0.25
H × VPA	-0.02 (0.20)	0.00
H × HV	0.06 (0.19)	0.64
H × OUTA	-0.11 (0.20)	0.57
HV × VPA	0.28 (0.19)	7.00†
HV × OUTA	0.11 (0.18)	0.48
L × VPA	0.13 (0.17)	1.01
L × H	0.08 (0.19)	0.58
L × OUTA	0.10 (0.20)	1.05
L × HV	0.20 (0.17)	2.01

I_{PSI} is an overall sexual isolation index for each pairwise comparison (range -1 to +1; null = 0, +1 = complete sexual isolation). Likelihood ratio values are from likelihood-ratio tests of the male population × female population interaction term in a logistic regression model, where a significant interaction term indicates sexual isolation. Using both analyses, stronger sexual isolation was detected for different-host comparisons than for same-host-comparisons (see also Fig. 1). $n = 64$ mating trials for each pairwise comparison. Population using *Adenostoma* are italicized; eight populations were used in total. Population codes and locations available from P.N.

* $P < 0.05$.

† $P < 0.01$.

sequent host adaptation, rather than genetic drift, has promoted the parallel evolution of sexual isolation. Visual predation has a crucial role in promoting sexual isolation among populations of *T. cristinae* by limiting gene flow between populations using different hosts. Three processes contribute to this reduction of gene flow. First, between-host migrants will be locally less cryptic and will exhibit reduced survivorship owing to predation^{5,6}. Second, mating success of between-host migrants will be reduced because *T. cristinae* mates assortatively by colour morph within populations and the less cryptic morph is rare (interpopulation mating experiment: $I_{PSI} = 0.46$, $n = 7$ pairs; additional laboratory matings: $I_{PSI} = 0.65$, $n = 128$ trials; copulating pairs collected in the wild: in *Ceanothus* populations, $I_{PSI} = 0.57$, $n = 88$, in *Adenostoma* populations, $I_{PSI} = 0.43$, $n = 64$; I_{PSI} measures sexual isolation between morphs within populations, all $P < 0.05$). Finally, even if locally less cryptic migrants secure mates, they will produce a higher proportion of locally less cryptic offspring because colour morph is genetically determined⁸. Given that host adaptation promotes sexual isolation and that gene flow erodes local adaptation⁶, differential predation accelerates the evolution of reproductive isolation by reducing the homogenizing effects of gene flow and facilitating adaptive differentiation through natural selection. □

Methods

Study species and location

Timema cristinae were collected from eight study sites in the Santa Ynez Mountains of California between February and April 2001. Individuals from different populations and the sexes were kept separate. Animals were fed foliage of the host on which they were collected and all the test animals used in mating experiments were juveniles that were reared to maturity in the laboratory.

Differentiation in morphology and cryptic resting behaviour

We conducted seven linear measurements on 550 walking-stick insects ($n = 259$ males,

291 females). Each trait was measured twice and measurement error was low for all traits (all repeatabilities greater than 0.90, $P < 0.001$, analysis of variance (ANOVA)). Mantel tests¹¹ were used to analyse associations between interpopulation distance matrices. The first principal component (PC1) from a PC analysis exhibited high and positive loadings for all traits in both sexes.

Individual walking-stick insects collected from both hosts ($n = 643$) were placed in the bottom of 500 ml plastic cups with one 12-cm host cutting from each host-plant species. These assays were initiated in the evening, and in the morning we recorded the position from which individuals were visible (for individuals choosing *Ceanothus* only, as the thin leaves of *Adenostoma* precluded measurement of this variable).

No-choice mating experiment

One male and one female were placed in a 10-cm Petri dish. At the end of one hour we scored whether the male and female were separate or copulating. We used walking-stick insects from eight populations, yielding 28 pairwise comparisons (13 same-host comparisons, 15 different-host comparisons). Sixteen sets of 16 individuals (one member of each sex from each population) were subjected to mating trials ($n = 64$ trials per pairwise comparison). Each individual was tested only once with an opposite-sex member from each of the eight populations, alternating the test order.

Statistical analysis of mating data

Copulation frequencies were analysed using logistic regression in a model that examined the dependence of copulation on male host, female host and an interaction term, with sexual isolation indicated by a significant interaction and significance assessed using a likelihood ratio test (LR). To test whether colour pattern contributed to sexual isolation, we used logistic regression to assess whether morph pair type (male and female the same or different colour-pattern morphs) in between-population crosses influenced the probability of copulation (using LR tests). To determine whether the degree of sexual isolation observed was dependent on which specific pair of populations was tested, we added male population and female population as terms in the previous logistic regression model and assessed the significance of the four-way interaction between all factors. We also tested for sexual isolation between populations for each pairwise comparison separately.

We calculated an index that summarizes the effects of sexual isolation (I_{PSI})¹², and used 5,000 re-samplings to calculate the standard deviation and significance of this index.

DNA sequencing

Individuals were collected from each of the populations used in the mating experiment (mtDNA (COI), $n = 40$, 5 individuals per population; nuclear DNA (ITS-2), $n = 23$, 2–4 individuals per population). For mtDNA, double-stranded PCR amplifications were performed with the primers S2183–A3014, S2195–A3014, S2183–A2887 and S2195–A2887 (ref. 13), which amplify most of the 3' half of the cytochrome oxidase I gene. For nuclear DNA, PCR amplifications were performed with the primers AED5.8F, 5'-TGTGAAGCTGAGGACACATGAAC-3' and 28SBLD, 5'-TTCTTTCTCTCC(C/G)CTTA(C/T)T(A/G)ATATGCTTAA-3'. Sequences are deposited in GenBank under accession numbers AF439805 to AF439820 (mtDNA) and AF459648 to AF459653 (nuclear DNA).

Phylogenetic analyses

Maximum likelihood trees were estimated using the HKY + G model of DNA substitution for mtDNA (COI) analyses and the F81 + G model for nuclear DNA (ITS-2) analyses, as these models gave the optimal fit to the data¹⁴. *Timema cristinae* from a different mountain (Ojala) was used as an outgroup for mtDNA analyses, whereas a closely related species (*T. monikensis*) was used as an outgroup for nuclear DNA analyses.

For mtDNA, analyses performed on unique haplotypes in PAUP 4.0b4b¹⁵ revealed that a molecular clock-constrained maximum likelihood tree ($-\ln$ likelihood = 915.07) did not differ significantly from a molecular clock-unconstrained tree with the same topology ($-\ln$ likelihood = 908.44; $\chi^2 = 13.26$, 14 d.f., $P > 0.50$)¹⁶. For nuclear DNA, a molecular clock was rejected (clock-unconstrained tree, $-\ln$ likelihood = 688.21; clock-constrained tree, $-\ln$ likelihood = 698.51; $\chi^2 = 0.60$, 6 d.f., $P < 0.05$).

The maximum likelihood mtDNA tree, showing bootstrap values from 200 replicates, was: (outgroup, (((P, L, H-HV-P): 82, L): 80, (P-VPC-VPA-OUTA-PR, (((H, H), (PR, PR-VPA-OUTA, VPC, VPC): 88): 64, (HV, H): 66): 90), (L, P): 94). The maximum likelihood nuclear DNA tree, showing bootstrap values from 200 replicates, was: (outgroup, P, (HV-L, (P-HV-L, L, (PR-VPC-OUTA-VPA-H, VPC): 54): 95): 73). Hyphens connect populations sharing a given haplotype, and colons precede bootstrap values. Populations using *Adenostoma* as a host are italicized. Eight allopatric populations were used.

Received 20 December 2001; accepted 5 March 2002.

1. Schluter, D. & Nagel, L. Parallel speciation by natural selection. *Am. Nat.* **146**, 292–301 (1995).
2. Rundle, H. D., Nagel, L., Boughman, J. W. & Schluter, D. Natural selection and parallel speciation in sticklebacks. *Science* **287**, 306–308 (2000).
3. Funk, D. J. Isolating a role for natural selection in speciation: host adaptation and sexual isolation in *Neochlamis bebbianae* leaf beetles. *Evolution* **52**, 1744–1759 (1998).
4. McPeck, M. A. & Wellborn, G. A. Genetic variation and reproductive isolation among phenotypically divergent amphipod populations. *Limnol. Oceanogr.* **43**, 1162–1169 (1998).
5. Sandoval, C. P. Differential visual predation on morphs of *Timema cristinae* (Phasmatodea: Timemidae) and its consequences for host range. *Biol. J. Linn. Soc.* **52**, 341–356 (1994).
6. Sandoval, C. P. The effects of relative geographic scales of gene flow and selection on morph frequencies in the walking stick *Timema cristinae*. *Evolution* **48**, 1866–1879 (1994).
7. Crespi, B. J. & Sandoval, C. P. Phylogenetic evidence for the evolution of ecological specialization in *Timema* walking-sticks. *J. Evol. Biol.* **13**, 249–262 (2000).
8. Sandoval, C. P. *Ecological and Behavioral Factors Affecting Spatial Variation in Colour or Morph Frequency in the Walking-stick Timema cristinae*. Thesis, Univ. California, Santa Barbara (1993).

9. Shimodaira, H. & Hasegawa, M. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol. Biol. Evol.* **16**, 1114–1116 (1999).
10. Templeton, A. R. Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and apes. *Evolution* **37**, 221–244 (1983).
11. Manly, B. F. J. *Randomization, Bootstrap and Monte Carlo Methods in Biology* (Chapman and Hall, London, 1997).
12. Rolán-Alvarez, E. & Caballero, A. Estimating sexual selection and sexual isolation effects from mating frequencies. *Evolution* **54**, 30–36 (2000).
13. Simon, C. *et al.* Evolution, weighting and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Ann. Ent. Soc. Am.* **87**, 651–701 (1994).
14. Posada, D. & Crandall, K. A. MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**, 817–818 (1998).
15. Swofford, D. L. *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods) 4.0b4a* (Sinauer Associates, Sunderland, Massachusetts, 2001).
16. Huelsenbeck, J. P. & Crandall, K. A. Phylogeny estimation and hypothesis testing using maximum likelihood. *Annu. Rev. Ecol. Syst.* **28**, 437–466 (1997).

Acknowledgements

We thank T. E. Reimchen, H. Rundle, A. Mooers, F. Breden, J. Endler, R. Vos, C. Parent, J. Joy, S. Springer and especially D. Schluter for discussion and comments on the manuscript. M. Fulton, D. McLaren, B. Mickelson, M. Vankoevinge and T. Luchin provided field and technical assistance. E. Rolán-Alvarez provided software and statistical advice. J. Endler provided laboratory space for all the mating experiments. Financial support was provided by the Natural Science and Engineering Research Council of Canada.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to P.N. (e-mail: pnosila@sfu.ca).

Netrin-1-mediated axon outgrowth requires deleted in colorectal cancer-dependent MAPK activation

Christelle Forcet^{*†‡}, Elke Stein^{†‡}, Laurent Pays^{*}, Véronique Corset^{*}, Fabien Llambi^{*}, Marc Tessier-Lavigne[‡] & Patrick Mehlen^{*}

^{*} Apoptosis/Differentiation Laboratory—label 'La Ligue'—Molecular and Cellular Genetic Center, CNRS UMR 5534, University of Lyon, 69622 Villeurbanne, France

[‡] Department of Biological Sciences, Howard Hughes Medical Institute, Stanford, California 94035-5020, USA

[†] These authors contributed equally to this work

Neuronal growth cones are guided to their targets by attractive and repulsive guidance cues¹. In mammals, netrin-1 is a bifunctional cue, attracting some axons and repelling others^{2–5}. Deleted in colorectal cancer (Dcc) is a receptor for netrin-1 that mediates its chemoattractive effect on commissural axons^{6,7}, but the signalling mechanisms that transduce this effect are poorly understood. Here we show that Dcc activates mitogen-activated protein kinase (MAPK) signalling, by means of extracellular signal-regulated kinase (ERK)-1 and -2, on netrin-1 binding in both transfected cells and commissural neurons. This activation is associated with recruitment of ERK-1/2 to a Dcc receptor complex. Inhibition of ERK-1/2 antagonizes netrin-dependent axon outgrowth and orientation. Thus, activation of MAPK signalling through Dcc contributes to netrin signalling in axon growth and guidance.

Signalling mechanisms triggered by netrin binding to Dcc are poorly understood. Dcc interacts with receptors of the Robo and UNC5H families, and adenosine A2b (refs 5, 8, 9), which are believed to modulate Dcc function. Beyond these cell-surface

interactions, one aspect of Dcc signalling already characterized is its role as a dependence receptor, inducing apoptosis in the absence of netrin-1 through a mechanism involving caspase cleavage of its cytoplasmic domain¹⁰. However, signalling events downstream of

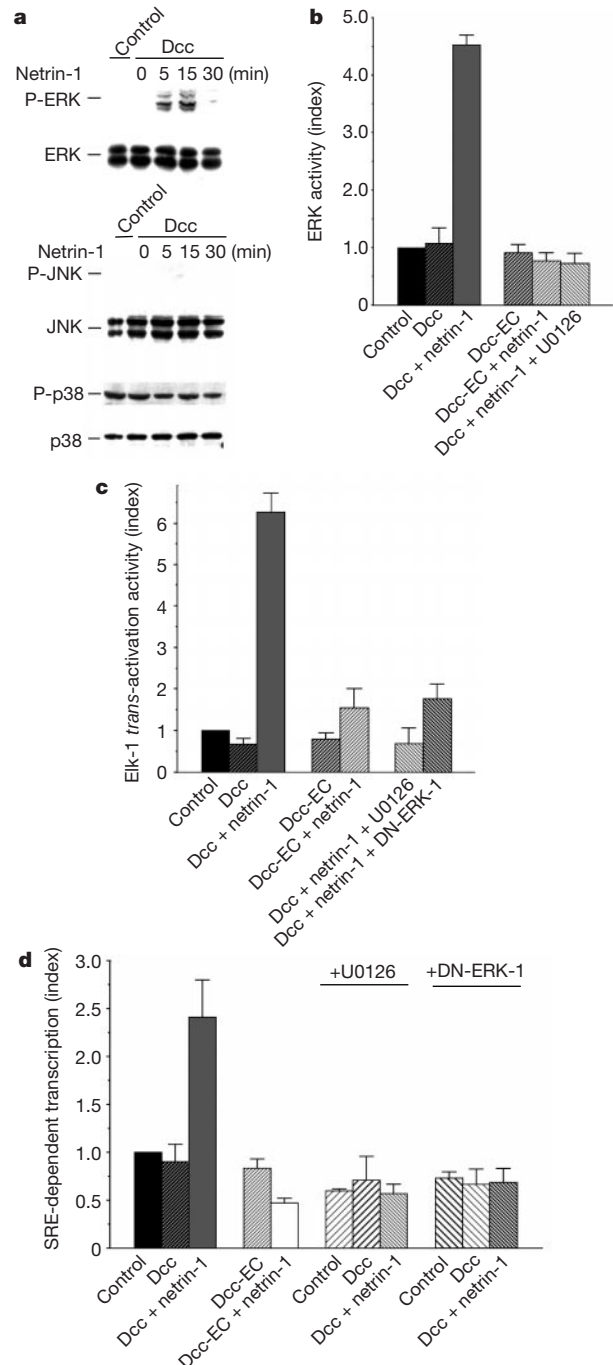


Figure 1 Dcc induces ERK-1/2 activation on netrin-1 binding. **a**, Western analysis of phosphorylated and total ERK-1/2, JNK and p38 in Dcc-transfected HEK 293 cells in response to 100 ng ml⁻¹ netrin-1. **b**, ERK kinase activity after 15 min of 100 ng ml⁻¹ netrin-1 in HEK 293 cells transiently expressing full-length or truncated Dcc (Dcc-EC; expression of the Dcc ectodomain). Control, 4 μM U0126. **c**, **d**, Elk-1 activation (**c**) or SRE-dependent gene expression (**d**) assessed by luciferase activity or SEAP expression, respectively, in HEK 293 cells transfected with full-length or truncated Dcc, with or without netrin-1 construct. Controls in **c** and **d**: co-transfection of dominant negative (DN) ERK-1, or addition of 4 μM U0126.

Dcc that mediate axon growth and guidance remain to be identified. We therefore examined whether classic signalling pathways mediate netrin actions.

The MAPK proteins, an evolutionarily conserved family of sequence- and structurally related enzymes, are interesting candidates for the downstream mediators of axon growth and guidance as they connect cell-surface receptors to critical regulatory targets¹¹. To test their involvement, HEK 293 or NIH3T3 cells were transiently transfected with a full-length Dcc expression construct, and phosphorylation of major MAPK proteins (ERK-1/2, Jun N-terminal kinase (JNK)-1/2 and p38) was analysed. Netrin-1 caused increased phosphorylation of ERK-1/2, but not JNK-1/2 or p38, in Dcc-transfected cells (Fig. 1a and not shown). This phosphorylation was rapid (appearing in less than 5 min) but transient, disappearing within 30 min (Fig. 1a). Increased phosphorylation was accompanied by increased ERK-1/2 activity, assessed by the ability of immunoprecipitated ERK-1/2 to phosphorylate myelin basic protein (MBP) *in vitro* (Fig. 1b and data not shown). This activation was not observed with a truncated form of Dcc that lacked its cytoplasmic domain. Activation was completely blocked by U0126, a specific inhibitor of MAP-kinase kinase (MEK)-1 and MEK-2 (Fig. 1b). Only slight inhibition, however, was observed with PD098059 (not shown), an inhibitor specific for MEK-1 over MEK-2 (ref. 12), suggesting that in these cell lines either MEK-2 mediates most of the netrin-induced ERK activation, or that either MEK-1 or -2 alone are redundant and must both be inhibited to block activation.

Some of the effects of ERK-1/2 are mediated through regulation of specific transcription factors¹³. We found that netrin activation of full-length, but not truncated, Dcc activated the transcription factor

Elk-1 (Fig. 1c). We also examined ERK-dependent transcription from a reporter gene controlled by a serum response element (SRE) enhancer element. Whereas Dcc alone had no effect, co-expression of Dcc and netrin-1 enhanced SRE-dependent transcription, an effect requiring the Dcc intracellular domain (Fig. 1d). Both Elk-1 activation and SRE-dependent transcription were blocked by U0126 or by expression of dominant negative ERK-1 (Fig. 1c, d). Thus, netrin-induced ERK-1/2 activation can cause transcriptional activation.

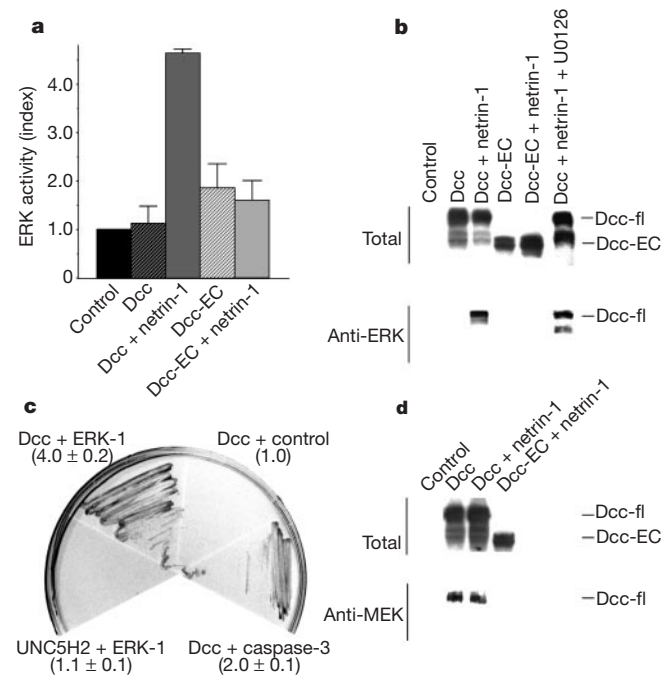


Figure 2 Interaction of Dcc with ERK-1/2 and MEK-1/2. HEK 293 cells transfected with mock, full-length (fl) or truncated Dcc-EC constructs were treated for 15 min with netrin-1 (**a**, **b**, **d**). **a**, ERK kinase assay performed on a Dcc immunoprecipitate. **b**, Immunoprecipitation using an anti-ERK-1/2 antibody followed by Dcc immunoblot. **c**, Interaction of the Dcc intracellular domain and ERK-1 assessed using the yeast two-hybrid system. Quantitative strength of the interaction is indicated in parentheses. **d**, Immunoprecipitation using a MEK-1/2 antibody followed by Dcc immunoblot. One Dcc molecule is immunoprecipitated by approximately nine ERK molecules or four MEK molecules (not shown).

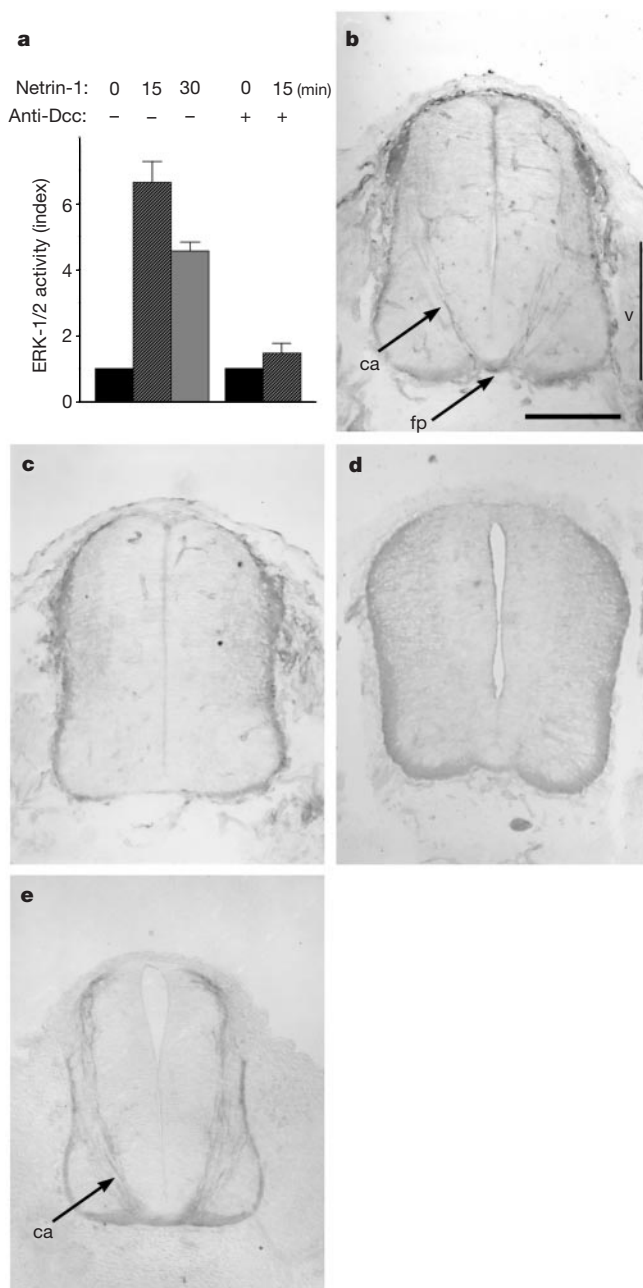


Figure 3 Netrin-1-induced ERK-1/2 activation occurs in commissural neurons. **a**, Rat E13 spinal cords were dissected and cultured in L15 medium⁴. A total of 100 ng ml⁻¹ netrin-1 was added for 0, 15 or 30 min in the presence or absence of 10 μg ml⁻¹ anti-Dcc (AF5) antibody, and ERK-1/2 activity was measured using an ERK kinase assay. **b–d**, ERK-1/2 activation occurs in commissural axons during spinal cord development. Phospho-ERK staining on transverse sections of spinal cords from E13 wild-type (**b**), *netrin-1*^{-/-} (**c**) or *Dcc*^{-/-} (**d**) embryos. **e**, Dcc staining. ca, commissural axons; fp, floor plate; v, ventral spinal cord. Scale bar, 200 μm.

To begin to investigate how Dcc activates ERK-1/2, we examined whether immunoprecipitates of Dcc and its interacting proteins contain ERK-1/2 activity. A 15 min netrin-1 treatment resulted in the presence of MBP phosphorylating activity and ERK-1/2 protein in a Dcc immunoprecipitate (Fig. 2a, b), but only from cells expressing full-length, not truncated, Dcc (Fig. 2a, b). In a control experiment, ERK-1/2 was activated by glial cell line derived neurotrophic factor (GDNF) binding to the Ret receptor tyrosine kinase expressed in HEK 293 cells, but did not immunoprecipitate Ret (not shown). In a yeast Gal4 two-hybrid assay, we found evidence for

direct interaction of the Dcc intracellular domain with full-length ERK-1 (Fig. 2c), suggesting that Dcc may directly recruit ERK-1 in response to netrin-1. This interaction seems to be required for netrin-induced ERK activation as it was abolished by deletion of a domain in Dcc required for the interaction (C.F., F.L., E.S., M.T.-L. and P.M., manuscript in preparation). As ERK activation usually occurs through formation of a module composed of a MAPK, a MAPKK and a MAPKKK protein¹¹, we tested whether the MAPKK MEK can interact with Dcc. Full-length, but not truncated, Dcc immunoprecipitated with endogenous MEK-1/2 (Fig. 2d); how-

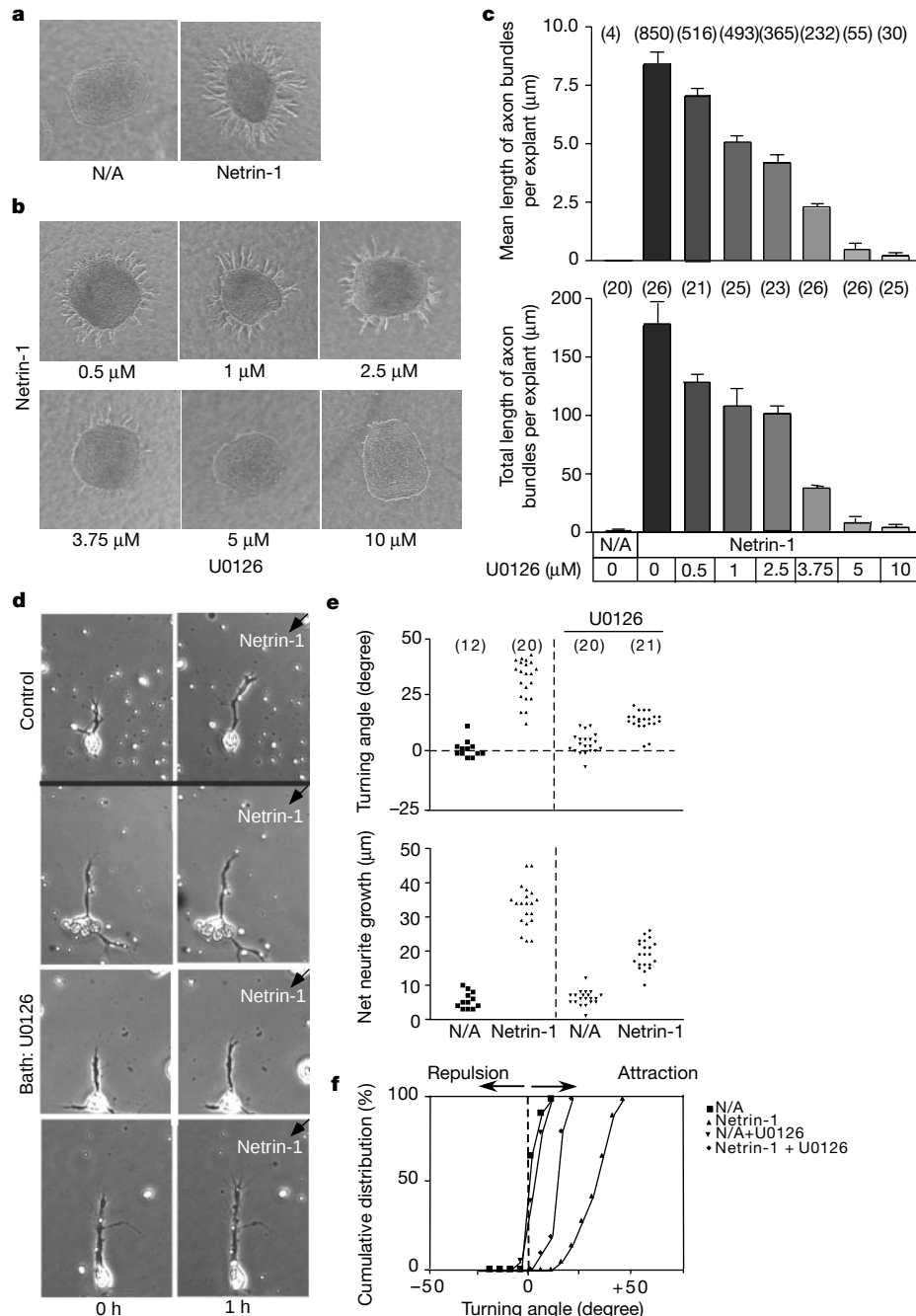


Figure 4 ERK-1/2 inhibition blocks netrin-mediated outgrowth and turning. **a**, E13 rat dorsal spinal cord explants cultured alone (N/A) or with netrin-1 (65 ng ml⁻¹). **b**, Explants cultured with netrin-1 (65 ng ml⁻¹) and increasing concentrations of U0126. **c**, Quantitative analysis. Numbers tested are in parentheses. **d**, Images of representative *Xenopus* spinal growth cones before and after a 1 h exposure to a netrin-1 gradient

(indicated by an arrow), with and without U0126 (10 μ M). **e**, Scatter diagram of turning angles (top) and net neurite extension during a 1-h period (bottom). Numbers of growth cones tested are in parentheses (symbols as in **f**). Cumulative distribution plot of turning angles (proportion with angles less than a given angle). Magnification: **a**, **b**, $\times 15$; **d**, $\times 300$.

ever, unlike ERK-1/2, MEK-1/2 interacted with Dcc equally in the presence or absence of netrin-1. These results suggest that MEK-1/2 may be latently bound to Dcc before recruitment and activation of ERK-1/2 in response to netrin-1; the recruitment, however, seems independent of MEK-1/2 activity, as it was not blocked by U0126 (Fig. 2b). Alternatively, we cannot exclude the possibility that MEK-1/2 bound nonspecifically to the intracellular domain of over-expressed Dcc. Future studies will determine whether this apparent docking mechanism triggers ERK-1/2 activation or instead simply amplifies activation stimulated by other factors.

We next examined whether ERK-1/2 contributes to the effects of Dcc on axon growth and guidance. We first tested whether ERK-1/2 activation occurs in a population of netrin-responsive neurons (commissural neurons). These neurons have cell bodies in the dorsal spinal cord that extend axons to the ventral spinal cord in response to netrin-1 (refs 2, 4). Netrin-1 stimulates outgrowth of commissural axons from explants of E13 (embryonic day 13) rat dorsal spinal cord by activating Dcc on these axons^{4,6}. Application of netrin-1 to isolated E13 rat spinal cord induced ERK-1/2 activation, with kinetics similar to those observed in Dcc-expressing HEK 293 cells (Fig. 3a). Dcc is required for this activation, as it was blocked by pre-treatment with a function-blocking anti-Dcc antibody⁶ (Fig. 3a). We next examined whether ERK-1/2 activation can occur specifically in commissural axons, using immunohistochemistry on tissue sections with an antibody to phosphorylated ERK-1/2. High levels of phospho-ERK-1/2 were detected in commissural axons between E10.5 and E13 (Fig. 3b and data not shown), ages at which the axons are growing to the floor plate². In netrin-1 and Dcc knockout animals we observed diminished phospho-ERK-1/2 staining in the ventral spinal cord (Fig. 3c, d); this finding is consistent with activation of ERK-1/2 by netrin-1, but its interpretation is complicated by the fact that commissural axon extension requires netrin-1 and Dcc, and is therefore impaired in these animals^{4,7}. Nonetheless, these results suggest that Dcc-dependent ERK activation occurs in commissural axons in response to netrin-1.

To determine whether ERK activation is required for netrin function, dorsal spinal cord explants from E13 rat embryos were grown for 16–18 h in collagen gels with or without netrin-1 (ref. 2). U0126 blocked netrin-1-induced axon extension in a dose-dependent manner ($P < 0.0001$ at all doses) (Fig. 4a–c). This effect does not represent a general inhibition of outgrowth but is, rather, specific for netrin-1 signalling, as the netrin-independent commissural axon outgrowth that can be observed when explants are grown for 40 h (ref. 6) was not affected by U0126 (total axon length per explant with and without U0126 (mean $\pm$ s.e.m.): $223 \pm 60 \mu\text{m}$ ($n = 18$) and $279 \pm 57 \mu\text{m}$ ($n = 17$), respectively ($P = 0.586$)). These observations support a requirement for ERK-1/2 in netrin-mediated outgrowth of commissural axons. In contrast to U0126, the p38 inhibitor SB-203580 actually increased the number of axon bundles emanating from explants in response to netrin-1 (not shown), an effect perhaps explained by a known side effect of SB-203580, that is, activation of c-Raf/ERK signalling¹⁴.

We next tested whether MAPK signalling is required for netrin-1 to re-orient axons, using an assay in which *Xenopus* spinal neurons turn towards a source of netrin-1; an effect mediated by Dcc^{5,15,16}. Whereas ERK-1/2 inhibition did not appear to affect the basal growth of axons, both netrin-1-induced turning and the netrin-1-stimulated increase in extension rate were antagonized by U0126 (Fig. 4d–f). Of note, the effects of netrin-1 were not completely blocked even though U0126 blocked ERK-1/2 activation by netrin-1 in *Xenopus* spinal neurons (not shown), suggesting an additional involvement of a MEK-1/2-independent pathway in the acute turning response.

Our results support a role for the MAPK pathway in responses to the chemoattractant netrin-1. This pathway has previously been implicated in neurite extension stimulated by neurotrophic factors

activating Trk receptors and by cell adhesion molecules like N-cadherin, laminin and L1 (refs 17–21). Conversely, activation of some repulsive receptors of the Eph family can repress MAPK signalling^{22,23}, although a causal role in repulsion has not been established. How does ERK activation affect axonal growth and guidance? Some effects probably result from phosphorylation of cytoskeletal ERK targets such as microtubule-associated proteins and neurofilaments²⁴. Interestingly, a MAPK-dependent mechanism drives internalization of the cell adhesion molecules ApCAM and L1 at the rear of the growth cone, which may allow protein cycling from the rear to the leading edge necessary for growth cone advance^{24–26}. We also show that Dcc-stimulated MAPK activation leads to activation of the transcription factor Elk-1 and SRE-regulated gene expression, providing a mechanism for transcriptional control by netrin-1. Elk-1 is present not just in neuronal cell bodies but also in axon terminals²⁷, but whether axonal Elk-1 participates in axon guidance is unknown. Other targets of ERK-1/2 involved in axon guidance may be translation regulators. New protein translation is stimulated by netrin-1 and required for netrin-mediated attraction of *Xenopus* retinal growth cones *in vitro*²⁸. Principal factors for translation initiation like eIF4E and eIF4E-BP1 are phosphorylated by an ERK-1/2-dependent pathway²⁹, providing a potential mechanism linking netrin-1 to protein synthesis for axon growth and guidance. □

Methods

Cell transfection, immunoblotting and netrin-1 production

Transient transfections of HEK 293 or NIH3T3 cells were performed using Lipofectamine following the manufacturer's suggested procedure (Life Technology). Immunoblots using different commercially available antibodies (phospho-ERK-1/2, Cell Signaling Technology; ERK-1/2, Sigma; phospho-p38, Cell Signaling Technology; p38, Cell Signaling Technology; phospho-JNK, Cell Signaling Technology; JNK, Cell Signaling Technology) were performed as described previously¹⁰. Netrin-1 was purified from netrin-1-producing HEK 293-EBNA cells².

Site-directed mutagenesis and plasmid constructs

The full-length Dcc-expressing construct pDcc-CMV-S and the netrin-1-expressing construct pGNET1-Myc were as described^{2,10}. The construct pDcc-EC-CMV, allowing the expression of the ectodomain of Dcc, was generated by a Quikchange strategy (Stratagene) replacing amino acid 1,121 by a stop codon. This was done using pDcc-CMV-S as template and the following primers: 5'-GCTGTGATTTCACCTGACTCGAGCGACGCTCTTCAGCC-3' and 5'-GGCTGAAGAGCGTCTCGAGTCAGGTGCAATACAGC-3'. The construct pRc/CMV-ERK1 T192A, coding for a dominant negative form of ERK-1, was provided by V. Fafeur. Haemagglutinin (HA)-tagged P44 (ERK-1) was provided by J. Pouyssegur. The two-hybrid construct pACT2-ERK-1 coding for the Gal4-activating domain fused to ERK-1 was generated by introducing, in the *Bam*H1-*Eco*R1 restriction sites of pACT2, the polymerase chain reaction product obtained from HA-P44 template and 5'-TATGGATCCGAGGCGGGGAGCCCCGG-3' and 5'-TATGAATCTTAGGGGCCTCTGG-3' primers.

MAPK activity assay

The MAPK immunoprecipitation kinase assay was performed using agarose-coupled ERK-1/2 antibody following Euromedex's instructions.

Elk-1 trans-reporting assay

HEK 293 or NIH3T3 cells were transiently transfected with the Dcc-expressing constructs together with pFA2-Elk-1 and pFR-Luc vectors (Stratagene). Elk-1 activation was measured using an Elk-1 trans-reporting assay (Stratagene) according to manufacturer's instructions.

Reporter gene assay

HEK 293 or NIH3T3 cells were transiently transfected with the Dcc- and/or netrin-1-expressing constructs together with pSRE-SEAP vector (Clontech). Twenty-four hours after transfection, SEAP was measured using Great EscAPE SEAP Detection System (Clontech).

Two-hybrid analysis

The two-hybrid system matchmaker II (Clontech) was performed as described previously⁸ using pAS2.1-Dcc-IC and pACT2-ERK-1. As a positive control, yeast cells were co-transformed with Dcc-IC and caspase-3 (ref. 8); as a negative control, yeast cells were co-transformed with Dcc and empty Gal4AD vector, or with UNC5H2 and ERK-1. The strength of the interaction was also measured following the manufacturer's instructions and is presented as the ratio between the level of β -galactosidase substrate degraded after 20 min for each co-transformed yeast and the one observed in the Dcc/Gal4AD control yeast.

Immunoprecipitation and immunohistochemistry

Immunoprecipitations were carried out on HEK 293 cells transfected with pDcc-CMV-5 or pDcc-EC-CMV using either anti-ERK-1/2 or anti-MEK-1/2 to immunoprecipitate endogenous ERK-1/2 or MEK-1/2, respectively. All immunoprecipitations were performed using 36 mg of lysates. The Dcc interaction with ERK-1/2 or MEK-1/2 was detected by immunoblot using an anti-Dcc antibody¹⁰.

We carried out production, breeding and genotyping of the netrin-1- and Dcc-deficient mice as described previously^{4,7}. Timed pregnant (plug date, E0) netrin-1 and Dcc mutant mice were killed, and individual embryos were fixed overnight at 4 °C using 4% paraformaldehyde in 0.1 M phosphate buffer, pH 7.4. Cryostat sections were stained with an anti-phospho ERK antibody (Cell Signaling Technology) or an anti-Dcc antibody. Immunostainings were visualized with a diaminobenzidine reaction.

Commissural axon outgrowth and *Xenopus* spinal turning assay

We performed both assays exactly as described¹⁶. The outgrowth from E13 explants was quantified by measuring the total length of all axon fascicles emerging from explants (total length of axon) or the mean of the length of these axon fascicles (mean length of axon bundles).

Received 26 February; accepted 18 April 2002.

Published online 1 May 2002, DOI 10.1038/nature748.

1. Tessier-Lavigne, M. & Goodman, C. S. The molecular biology of axon guidance. *Science* **274**, 1123–1133 (1996).
2. Serafini, T. *et al.* The netrins define a family of axon outgrowth-promoting proteins homologous to *C. elegans* UNC-6. *Cell* **78**, 409–424 (1994).
3. Kennedy, T. E., Serafini, T., de la Torre, J. R. & Tessier-Lavigne, M. Netrins are diffusible chemotropic factors for commissural axons in the embryonic spinal cord. *Cell* **78**, 425–435 (1994).
4. Serafini, T. *et al.* Netrin-1 is required for commissural axon guidance in the developing vertebrate nervous system. *Cell* **87**, 1001–1014 (1996).
5. Hong, K. *et al.* A ligand-gated association between cytoplasmic domains of UNC5 and DCC family receptors converts netrin-induced growth cone attraction to repulsion. *Cell* **97**, 927–941 (1999).
6. Keino-Masu, K. *et al.* Deleted in Colorectal Cancer (DCC) encodes a netrin receptor. *Cell* **87**, 175–185 (1996).
7. Fazeli, A. *et al.* Phenotype of mice lacking functional *Deleted in colorectal cancer (Dcc)* gene. *Nature* **386**, 796–804 (1997).
8. Corset, V. *et al.* Netrin-1-mediated axon outgrowth and cAMP production requires interaction with adenosine A2b receptor. *Nature* **407**, 747–750 (2000).
9. Stein, E. & Tessier-Lavigne, M. Hierarchical organization of guidance receptors: silencing of netrin attraction by slit through a Robo/DCC receptor complex. *Science* **291**, 1928–1938 (2001).
10. Mehlen, P. *et al.* The DCC gene product induces apoptosis by a mechanism requiring receptor proteolysis. *Nature* **395**, 801–804 (1998).
11. Chang, L. & Karin, M. Mammalian MAP kinase signalling cascades. *Nature* **410**, 37–40 (2001).
12. Alessi, D. R., Cuenda, A., Cohen, P., Dudley, D. T. & Saltiel, A. R. PD 098059 is a specific inhibitor of the activation of mitogen-activated protein kinase kinase *in vitro* and *in vivo*. *J. Biol. Chem.* **270**, 27489–27494 (1995).
13. Garrington, T. P. & Johnson, G. L. Organization and regulation of mitogen-activated protein kinase signalling pathways. *Curr. Opin. Cell Biol.* **11**, 211–218 (1999).
14. Birkenkamp, K. U. *et al.* p38MAP kinase inhibitor SB203580 enhances nuclear factor- κ B transcriptional activity by a non-specific effect upon the ERK pathway. *Br. J. Pharmacol.* **131**, 99–107 (2000).
15. de la Torre, J. R. *et al.* Turning of retinal growth cones in a netrin-1 gradient mediated by the netrin receptor DCC. *Neuron* **19**, 1211–1224 (1997).
16. Stein, E., Zou, Y., Poo, M.-M. & Tessier-Lavigne, M. Binding of DCC by netrin-1 to mediate axon guidance independent of adenosine A2b receptor activation. *Science* **291**, 1976–1982 (2001).
17. Fukuda, M. *et al.* Induction of neurite outgrowth by MAP kinase in PC12 cells. *Oncogene* **11**, 239–244 (1995).
18. Kolkova, K., Novitskaya, V., Pedersen, N., Berezin, V. & Bock, E. Neural cell adhesion molecule-stimulated neurite outgrowth depends on activation of protein kinase C and the Ras-mitogen-activated protein kinase pathway. *J. Neurosci.* **20**, 2238–2246 (2000).
19. Schmid, R. S. *et al.* NCAM stimulates the Ras-MAPK pathway and CREB phosphorylation in neuronal cells. *J. Neurobiol.* **38**, 542–558 (1999).
20. Schmid, R. S., Pruitt, W. M. & Maness, P. F. A MAP kinase-signalling pathway mediates neurite outgrowth on L1 and requires Src-dependent endocytosis. *J. Neurosci.* **20**, 4177–4188 (2000).
21. Yamashita, H., Avraham, S., Jiang, S., Dikic, I. & Avraham, H. The Csk homologous kinase associates with TrkA receptors and is involved in neurite outgrowth of PC12 cells. *J. Biol. Chem.* **274**, 15059–15065 (1999).
22. Elowe, S., Holland, S. J., Kulkarni, S. & Pawson, T. Downregulation of the Ras-mitogen-activated protein kinase pathway by the EphB2 receptor tyrosine kinase is required for ephrin-induced neurite retraction. *Mol. Cell Biol.* **21**, 7429–7441 (2001).
23. Miao, H. *et al.* Activation of EphA receptor tyrosine kinase inhibits the Ras/MAPK pathway. *Nature Cell Biol.* **3**, 527–530 (2001).
24. Gundersen, G. G. & Cook, T. A. Microtubules and signal transduction. *Curr. Opin. Cell Biol.* **11**, 81–94 (1999).
25. Schaefer, A. W. *et al.* Activation of the MAPK signal cascade by the neural cell adhesion molecule L1 requires L1 internalization. *J. Biol. Chem.* **274**, 37965–37973 (1999).
26. Kamiguchi, H. & Lemmon, V. Recycling of the cell adhesion molecule L1 in axonal growth cones. *J. Neurosci.* **20**, 3676–3686 (2000).
27. Sgambato, V. *et al.* *In vivo* expression and regulation of Elk-1, a target of the extracellular-regulated kinase signalling pathway, in the adult rat brain. *J. Neurosci.* **18**, 214–226 (1998).
28. Campbell, D. S. & Holt, C. E. Chemotropic responses of retinal growth cones mediated by rapid local protein synthesis and degradation. *Neuron* **32**, 1013–1026 (2001).
29. Herbert, T. P., Tee, A. R. & Proud, C. G. The extracellular signal-regulated kinase pathway regulates the phosphorylation of 4E-BP1 at multiple sites. *J. Biol. Chem.* **277**, 11591–11596 (2002).

Acknowledgements

We thank J. Pouyssegur for reagents and discussions; V. Fafeur and E. Bloch-Gallego for providing reagents; C. Guix for technical assistance; and S. Dazy for preliminary observations on ERK-1/2 activation. This work was supported by the Ligue Contre le Cancer and the ARC (P.M.). L.P. is supported by a post-doctoral fellowship from the Ligue Contre le Cancer. E.S. is an Associate and M.T.-L. an Investigator of the Howard Hughes Medical Institute.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to P.M. (e-mail: mehlen@univ-lyon1.fr).

Molecular identification of a renal urate–anion exchanger that regulates blood urate levels

Atsushi Enomoto*†, Hiroaki Kimura*‡, Arthit Chairoungdua*, Yasuhiro Shigeta*§, Promsuk Jutabha*, Seok Ho Cha*, Makoto Hosoyamada*, Michio Takeda*, Takashi Sekine||, Takashi Igarashi||, Hirotaka Matsuo*¶, Yuichi Kikuchi#, Takashi Oda☆, Kimiyoshi Ichida‡, Tatsuo Hosoya‡, Kaoru Shimokata†, Toshimitsu Niwa†, Yoshikatsu Kanai* & Hitoshi Endou*

* Department of Pharmacology and Toxicology, Kyorin University School of Medicine, Tokyo 181-8611, Japan

† Department of Clinical Preventive Medicine, Nagoya University School of Medicine, Nagoya 466-8560, Japan

‡ Division of Kidney and Hypertension, Department of Internal Medicine, The Jikei University School of Medicine, Tokyo 105-8461, Japan

§ Department of Urology, Chiba University School of Medicine, Chiba 260-8670, Japan

|| Department of Pediatrics, Faculty of Medicine, The University of Tokyo, Tokyo 113-8655, Japan

¶ First Department of Physiology and # Second Department of Internal Medicine, National Defense Medical College, Saitama 359-8513, Japan

☆ Department of Medicine, JSDF Kumamoto Hospital, Kumamoto 862-0902, Japan

Urate, a naturally occurring product of purine metabolism, is a scavenger of biological oxidants implicated in numerous disease processes^{1–3}, as demonstrated by its capacity of neuroprotection^{4,5}. It is present at higher levels in human blood (200–500 μ M) than in other mammals⁶, because humans have an effective renal urate reabsorption system, despite their evolutionary loss of hepatic uricase by mutational silencing^{6–8}. The molecular basis for urate handling in the human kidney remains unclear because of difficulties in understanding diverse urate transport systems and species differences^{6,9,10}. Here we identify the long-hypothesized^{9–11} urate transporter in the human kidney (URAT1, encoded by *SLC22A12*), a urate–anion exchanger regulating blood urate levels and targeted by uricosuric and anti-uricosuric agents (which affect excretion of uric acid). Moreover, we provide evidence that patients with idiopathic renal hypouricaemia (lack of blood uric acid) have defects in *SLC22A12*. Identification of URAT1 should provide insights into the nature of urate homeostasis, as well as lead to the development of better agents against hyperuricaemia, a disadvantage concomitant with human evolution.

Urate exists primarily as a weak acid at physiological pH with a pK_a value of 5.75 (ref. 10); therefore, the structure of the urate transporter should be similar to that of members of the organic

anion transporter (OAT) family^{10,12–14}. We systematically identified paralogues of OATs in the entire draft human genome database¹⁵, and identified a sequence that seemed to be similar to the *OAT4* gene¹⁴ in chromosome 11q13. We predicted the structure of this gene (*SLC22A12*), and succeeded in isolating the complementary DNA (URAT1, 2,642 base pairs, bp) from human kidney poly(A)⁺ RNA by 3' rapid amplification of cDNA ends (3'-RACE). The amino-acid sequence predicted by the URAT1 cDNA exhibits 42% amino-acid identity with that of *OAT4* (ref. 14) (Fig. 1a). The hydropathy plot predicted 12 membrane-spanning domains in URAT1, which are similar to those in members of OAT family^{12–14}. Northern blotting of human tissues revealed substantial hybridization with a 2.8-kilobase (kb) transcript in human adult and fetal kidney (Fig. 1b).

Immunohistochemical analysis of human kidney revealed that URAT1 protein was prominent in epithelial cells of proximal tubules in the renal cortex (Fig. 1c). Under high magnification, the URAT1 protein was found to be located in the luminal membrane of the epithelium of proximal tubules (Fig. 1d) but not in that of distal tubules. In western blot analysis of human kidney protein (Fig. 1e), the anti-URAT1 antibody recognized a band of relative molecular mass 40,000 (*M_r* 40K) that was found to be affected by the presence of an antigen peptide (50 µg ml⁻¹) in the absorption experiment.

Table 1 Substrate specificity of URAT1

	Uptake (% of urate uptake with no inhibitor)
Water injected	14.1 ± 1.8
Control (no inhibitor)	100.0 ± 15.2
Urate (1 mM)	18.3 ± 1.5**
L-Lactate (1 mM)	64.6 ± 6.3*
L-Lactate (10 mM)	46.4 ± 8.0**
D-Lactate (1 mM)	71.1 ± 11.2
D-Lactate (10 mM)	25.5 ± 2.12**
Nicotinate (1 mM)	26.1 ± 3.2**
Probenecid (1 mM)	19.1 ± 3.9**
Benzbromarone (50 µM)	6.9 ± 2.4**
Pyrazinecarboxylic acid (1 mM)	27.7 ± 3.1**
Pyrazinamide (1 mM)	126.1 ± 10.2
Pyrazine (1 mM)	120.0 ± 11.5
Allopurinol (1 mM)	89.9 ± 6.8
Xanthine (1 mM)	104.1 ± 18.3
Pyrazinedicarboxylic acid (1 mM)	96.1 ± 3.0
Pyrazinedicarboxylic acid (10 mM)	39.2 ± 5.1**
Acetoacetate (1 mM)	98.0 ± 5.1
Acetoacetate (10 mM)	51.3 ± 5.8**
Succinate (1 mM)	89.3 ± 11.2
Succinate (10 mM)	23.7 ± 5.2**
Ketoglutarate (1 mM)	97.2 ± 12.3
Ketoglutarate (10 mM)	73.4 ± 5.1*
β-hydroxybutyric acid (1 mM)	80.5 ± 4.9
β-hydroxybutyric acid (10 mM)	50.0 ± 4.3**
Orotic acid (1 mM)	26.5 ± 4.9**
Phenylbutazone (1 mM)	29.1 ± 3.2**
Sulfinpyrazone (1 mM)	39.6 ± 7.7**
Furosemide (1 mM)	38.4 ± 10.3**
Bumetanide (1 mM)	26.2 ± 4.5**
PAH (1 mM)	91.5 ± 7.9
Estrone sulphate (1 mM)	126.1 ± 16.4
Salicylic acid (1 mM)	21.9 ± 3.4**
Indomethacin (1 mM)	8.9 ± 2.6**
Methotrexate (1 mM)	86.8 ± 15.2
TEA (1 mM)	92.5 ± 7.5
Choline chloride (1 mM)	86.5 ± 13.5
DIDS (200 µM)	54.3 ± 2.1**
DIDS (1 mM)	34.1 ± 2.2**
NPPB (200 µM)	20.1 ± 3.6**
Losartan (1 mM)	15.9 ± 6.6**
EXP-3174 (1 mM)	23.1 ± 2.2**

The uptake rate of [¹⁴C]urate (50 µM) by URAT1-expressing oocytes occurring in exchange for Cl⁻ in the 0 Cl⁻ bath was determined at 1 h in the absence or presence of inhibitors, which were added to the extracellular medium (pH 7.4) at the indicated concentrations. The values were expressed as percentages of [¹⁴C]urate uptake by URAT1-expressing oocytes under control conditions (absence of the inhibitors). TEA, tetraethylammonium; DIDS, 4,4'-diisothiocyanatostilbene-2,2'-disulphonic acid; NPPB, 5-nitro-2-(3-phenylpropylamino) benzoic acid. Data are mean ± s.e.m with *n* = 8–10. *, *P* < 0.05; **, *P* < 0.001, compared with the uptake in the presence of no inhibitors.

Xenopus oocytes injected with URAT1 complementary RNA exhibited time-dependent transport activity of [¹⁴C]urate (Fig. 2a) but not of various typical substrates of OATs^{10,12–14} or organic cation transporters (OCTs)^{16,17}. The URAT1-mediated urate uptake, a carrier-mediated transport process, could be saturated. The esti-

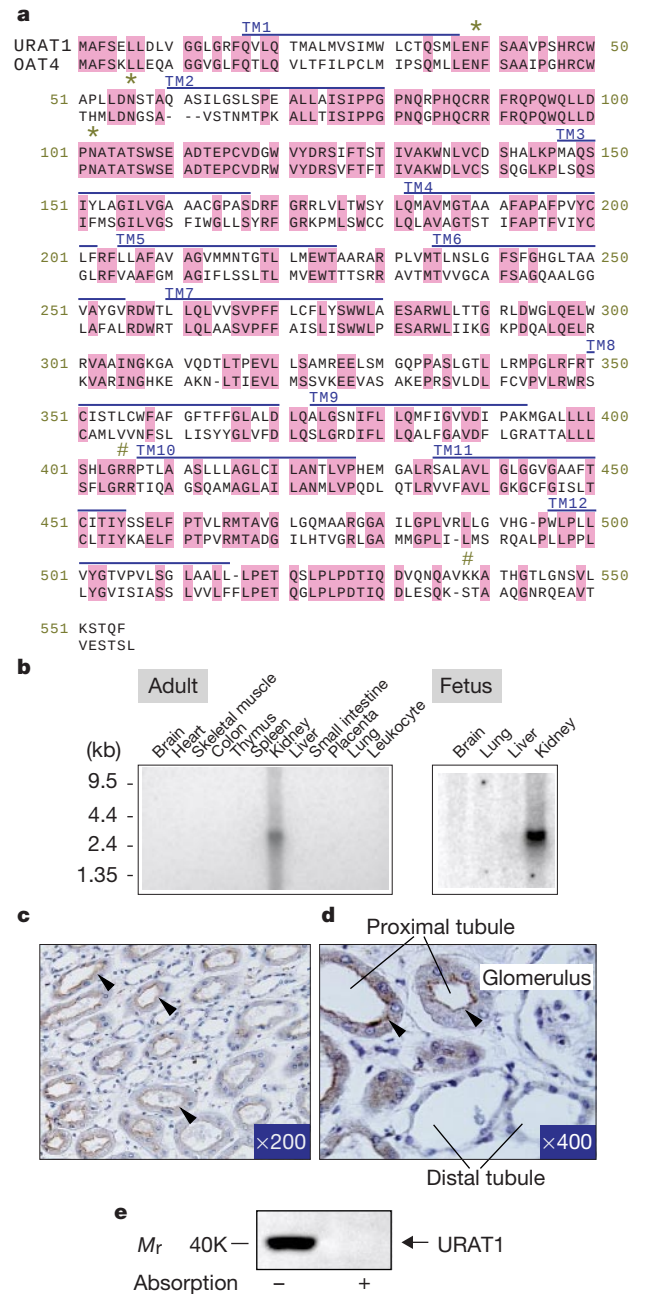


Figure 1 Sequence analysis and localization of URAT1. **a**, The amino-acid sequence of URAT1 (encoded by *SLC22A12*) is aligned with that of *OAT4* (ref. 14) (encoded by *SLC22A11*). URAT1 has three consensus sites for *N*-glycosylation (asterisks: N39, N56, N102) and two cAMP-dependent protein kinase phosphorylation sites (hash signs: 405RRPT, 536KKAT). Lines above the sequence (TM1–TM12) indicate the *trans*-membrane domains in URAT1. Identical sequences between URAT1 and *OAT4* are shaded. **b**, High-stringency northern analysis revealed predominant expression of URAT1 mRNA in the human adult (left) and fetal (right) kidney. kb, kilobases. **c**, **d**, Immunohistochemical detection of the URAT1 protein in the human kidney. **e**, Western blot analysis of the human renal protein under reducing conditions. The absorption test was performed by preincubation of the URAT1 peptide (50 µg ml⁻¹) with the antibody.

mated Michaelis constant (K_m) for urate was $371 \pm 28 \mu\text{M}$ ($n = 3$) (Fig. 2b). These data, together with the luminal location of URAT1 in proximal epithelial cells, allow us to propose a model that the URAT1-mediated pathway constitutes a mechanism for reabsorption of urate from the tubular lumen (extracellular) to the cytosol (intracellular) at the proximal tubules. The existence of a urate transporter has been suggested in an investigation using brush-border membrane vesicles (BBMV) from humans¹⁸.

The mode of urate transport via URAT1 was revealed by substitution experiments with Na^+ and K^+ , which demonstrated that it does not involve a direct Na^+ -urate cotransport and is not driven by membrane potential (Fig. 2c). We then measured the effect of an outward Cl^- gradient (intracellular to extracellular) on urate uptake. Imposing a Cl^- gradient by complete removal of external Cl^- markedly accelerated the urate uptake via URAT1 (Fig. 2c). We investigated Cl^- efflux from oocytes preloaded with Na^{36}Cl (Fig. 2d). The extent of stimulation of Cl^- efflux by extracellular urate was statistically significant, demonstrating the existence of an exchange mechanism for urate and inorganic Cl^- in URAT1. Other halide anions can be substituted for Cl^- in exchange for urate in URAT1 (Fig. 2e). Imposing a Br^- or I^- gradient, but not a F^- gradient, markedly accelerated urate uptake compared with imposing a Cl^- gradient, indicating that Br^- and I^- but not F^- can be substituted for Cl^- and coupled to URAT1-mediated urate transport. The dependence on extracellular or intracellular pH was not observed in urate transport, which is consistent with previous investigations that the urate transporter in human renal BBMV¹⁸ does not have an affinity for OH^- , in contrast to urate transporters in other animals^{9,10,19,20} (see Supplementary Information).

URAT1 exhibited pharmacological properties consistent with those of previously described urate transport activity in human BBMV^{10,18} (Table 1). We found that the urate transport via URAT1 is inhibited selectively by organic anions such as lactate, nicotinate, acetoacetate, hydroxybutyrate and succinate. Aromatic monocarboxylates such as pyrazinecarboxylic acid (PZA), nicotinate and orotate at 1 mM inhibited urate uptake more effectively than aliphatic monocarboxylates such as lactate. Para-aminohippurate (PAH), the representative substrate of OATs^{10,12,13}, did not exert an inhibitory effect on urate uptake via URAT1, consistent with the observation that PAH has no effect on the fractional excretion of urate in humans¹⁸. We also tested the influence of uricosuric substances known to reduce hyperuricaemia^{10,21}. Probenecid, phenylbutazone, sulphapyrazole, nonsteroidal anti-inflammatory drugs, and diuretic drugs at 1 mM *cis* inhibited urate uptake; benzbromarone was the most potent, completely inhibiting urate uptake at $50 \mu\text{M}$, with a 50%-inhibitory concentration (IC_{50}) of less than $0.1 \mu\text{M}$ (see Supplementary Information). Losartan, an antagonist of the angiotensin II receptor ($\text{AT}_1\text{-RA}$), and its metabolite (EXP-3174) *cis* inhibited urate uptake (Table 1) and the IC_{50} of losartan tended to be the same as that of probenecid (see Supplementary Information), consistent with the uricosuric property of losartan^{9,10}.

To test whether the compounds that inhibited URAT1-mediated urate uptake are transported via URAT1, we performed *trans*-stimulation experiments. Intracellularly loaded lactate, PZA and nicotinate strongly stimulated urate uptake (Fig. 3a); PZA, an active metabolite of pyrazinamide widely used as an antituberculosis agent, was most potent among these anions, which is consistent with its capacity to alter renal urate clearance^{10,22}. As expected from the lack of *cis* inhibition on urate uptake by PAH, intracellularly loaded PAH had no *trans*-stimulatory effect. The *trans*-stimulatory effects of preloaded NaCl or KCl was statistically significant, but the effects were small compared with that of PZA. Therefore, the major counteranions that exchange for urate via URAT1 are organic anions, rather than inorganic Cl^- . Next, we tested the *trans*-stimulatory effect on urate uptake by other anions relevant to the proximal tubule, for example, bicarbonate, nitrate and sulphate.

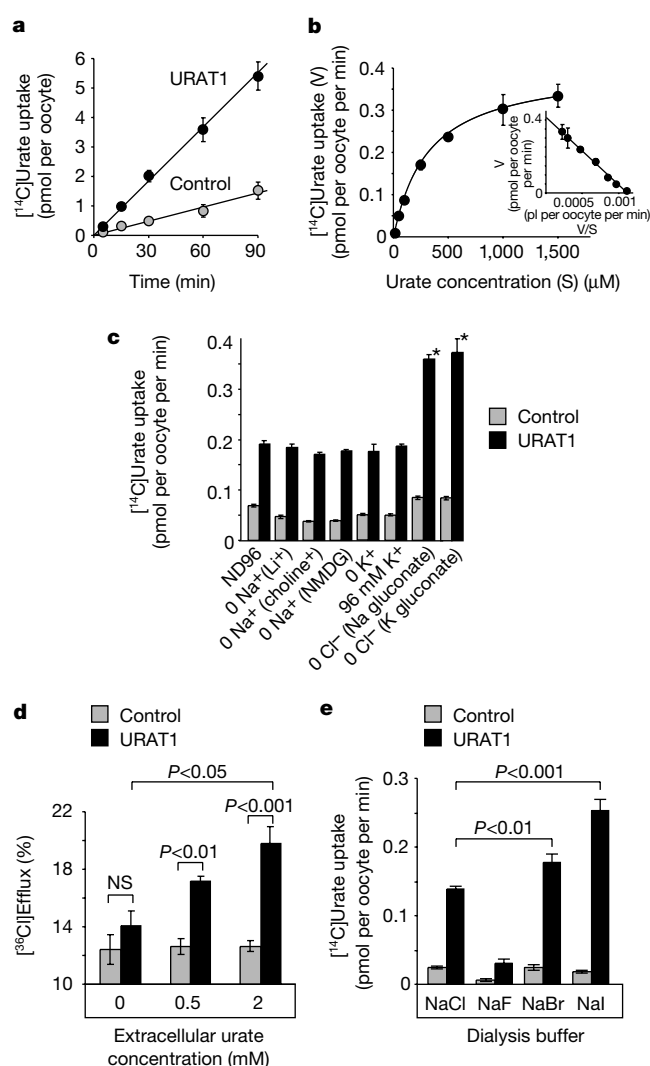


Figure 2 Functional expression of URAT1 in *Xenopus* oocytes. **a**, Time course of urate uptake by URAT1. The uptake of $50 \mu\text{M}$ ^{14}C urate in water-injected oocytes (control) and URAT1-expressing oocytes was measured during a 90-min incubation. **b**, Transport kinetics of URAT1. Urate transport was measured in control oocytes and URAT1-expressing oocytes over a urate concentration range of 10–1,500 μM . The URAT1-specific transport activity in URAT1-expressing oocytes was calculated by subtracting the transport activity in control oocytes. These URAT1-specific transport activities were used in kinetic analysis. Inset, Eadie–Hofstee plot. **c**, Urate uptake by URAT1 is stimulated by Cl^- removal from the bath. The transport rate of ^{14}C urate ($50 \mu\text{M}$) in control oocytes and URAT1-expressing oocytes was measured for 1 h in the presence or absence of extracellular Na^+ , K^+ and Cl^- . In 0 Na^+ bath, Na^+ in the ND96 bath was replaced with Li^+ , choline or *N*-methyl-D-glucamine (NMDG). In 96 mM K^+ bath, Na^+ was replaced with K^+ . In 0 Cl^- bath, NaCl in the ND96 bath was replaced with Na gluconate or K gluconate. Gluconate, which replaced Cl^- , did not exert a *trans*-stimulatory effect on urate transport as described in Fig. 3a. Asterisk, $P < 0.001$ compared with ND96. **d**, *Trans*-stimulatory effect of extracellular urate on URAT1-mediated efflux of Cl^- . URAT1-expressing and control oocytes were injected with 50 nl of Na^{36}Cl ($\sim 8.89 \text{ mCi ml}^{-1}$) using fine-tipped micropipettes. Individual oocytes were washed and transferred to the 0 Cl^- solution or 0.5 or 2 mM unlabelled urate. The amount of Cl^- efflux for 90 min is shown as the percentage of the preloaded amount. NS, not significant. **e**, Other halide anions can be substituted for Cl^- in exchange for urate in URAT1. Intracellular Cl^- was depleted and substituted for other halide anions by dialysis of URAT1-expressing and control oocytes for 16–24 h at 16°C in a Cl^- -free medium (Cl^- in ND96 was replaced by F^- , Br^- or I^-); these oocytes were compared with matched oocytes incubated in ND96. The transport rate of ^{14}C urate ($50 \mu\text{M}$) in these oocytes was measured for 1 h in the 0 Cl^- bath. All data are mean $\pm$ s.e.m. with $n = 8$ –10.

Among these anions, only sodium nitrate (NaNO_3) exerted a *trans*-stimulatory effect on urate uptake with statistical significance (Fig. 3a). We investigated the efflux of nicotinate from control or URAT1-expressing oocytes preloaded with [^3H]nicotinate (Fig. 3b). Extracellular urate (2 mM) significantly stimulated the URAT1-mediated efflux of nicotinate. Thus, it is suggested that an exchange mechanism for urate and these organic anions is present in URAT1.

As illustrated in Fig. 3c, intracellular accumulation of those organic anions for which URAT1 has affinity will favour the uphill reabsorption of urate in exchange for these anions, which move down their electrochemical gradients into the tubular lumen. Intracellular anions are made available by their uptake from the glomerular filtrate across the luminal membrane²³, by uptake from peritubular capillaries by basolateral transporters (OATs)^{10,12,13}, and by cell metabolism²⁴. Drugs or agents with affinity for URAT1 will be uricosuric when acting from the lumen, whereas

they will be antiuricosuric by driving the urate influx when acting from the intracellular space, consequently regulating blood urate levels.

Patients with idiopathic renal hypouricaemia (MIM number 220150) with exercise-induced acute renal failure and chronic renal dysfunction have previously been described^{25,26}. One of these patients (a 48-year-old male) had a blood urate level of $4.3 \pm 0.9 \mu\text{g ml}^{-1}$ (normal range is $55\text{--}70 \mu\text{g ml}^{-1}$). The fractional excretion of urate was $95\% \pm 10$, (normally $<10\%$), and the patient showed only mild responses in pyrazinamide and benzbromarone loading tests (ref. 25), indicating that he had a selective defect in the urate reabsorption mechanism of his kidneys. Analysis of polymerase chain reaction (PCR) products from the URAT1-coding region from the patient's genomic DNA revealed the presence of a homozygous G $\rightarrow$ A transition at nucleotide 774 within exon 4 of *SLC22A12* (Fig. 4a), changing tryptophan 258 (TGG) to a stop codon (TGA) and producing a truncated protein that lacked half of the mature protein (the structure of *SLC22A12* is included in Supplementary Information). We also detected two homozygous missense mutations—650 C $\rightarrow$ T (in exon 3) and 894 G $\rightarrow$ T (in exon 5)—which lead to, respectively, the substitution of threonine by methionine (T217M) and glutamic acid by aspartic acid (E298D) in unrelated individuals affected with renal hypouricaemia (data not shown). None of these homozygous mutations were identified in 180 randomly chosen control Japanese individuals. We analysed urate transport after injection of wild-type or mutated cRNA into oocytes and found that these substitutions abolished urate transport activity without altering their expression or membrane transport (Fig. 4b, c). This finding reveals a pivotal role of URAT1 in the accumulation of much higher urate levels in humans than in other mammals. The fractional excretion of urate is normally less than 10% in humans, in contrast to other species in which the urate excretion rate exceeds the rate at which it is being filtered^{9,10}. It would be interesting to determine whether urate nephrolithiasis²⁷ is also associated with defects in *SLC22A12*. In addition, although only speculative, some common polymorphisms in *SLC22A12* may be

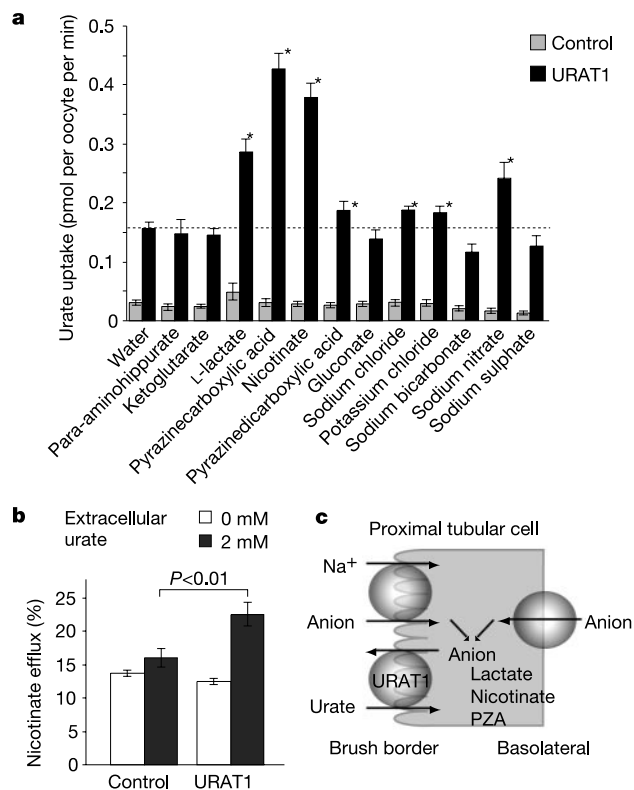


Figure 3 Urate uptake via URAT1 is *trans*-stimulated by intracellular anions. **a**, *Trans*-stimulatory effect of preloaded organic or inorganic anions on URAT1-mediated urate uptake. Control or URAT1-expressing oocytes were injected with 50 nl of water or unlabelled test anions (at 100 mM, pH 7.4) using fine-tipped micropipettes and then transferred to the 0 Cl^- bath containing $50 \mu\text{M}$ [^{14}C]urate; the uptake rate for 1 h was determined. Asterisk, $p < 0.05$ compared with water. **b**, *Trans*-stimulatory effect of extracellular urate on URAT1-mediated efflux of nicotinate. Control or URAT1-expressing oocytes were injected with 50 nl of [^3H]nicotinate ($\sim 50\text{--}60 \text{ Ci mmol}^{-1}$), transferred to the 0 Cl^- solution with or without unlabelled urate (2 mM), and incubated at room temperature. The amount of [^3H]nicotinate efflux for 90 min is shown as the percentage of the preloaded amount. All data are mean $\pm$ s.e.m. with $n = 8\text{--}10$. **c**, Proposed model of indirect coupling of sodium and urate transport via URAT1. Anions are coupled to sodium uptake along the luminal membrane and are later exchanged for urate by URAT1 in the proximal tubule. Organic anions that are actively pumped into proximal cells from apical or basolateral sides or those that are produced in cells should favour urate reabsorption, by leaving the cells in exchange for luminal urate. A clone responsible for the electrogenic Na^+ -anion cotransporter has not yet been identified, although a Na^+ -dicarboxylate cotransporter localized to the S_3 segment has been identified (NaDC-1)¹⁰.

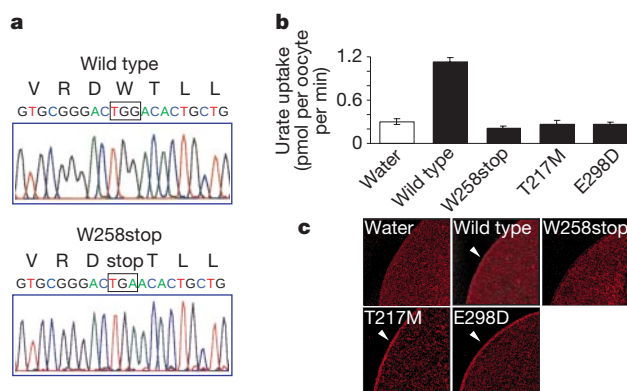


Figure 4 Mutations within *SLC22A12* are associated with idiopathic renal hypouricaemia. **a**, Electropherograms of partial sequences of exon 4 of *SLC22A12* showing the homozygous point mutation (W258stop) found in a patient (a 48-year-old male). This homozygous W258stop mutation was harboured in two other patients exhibiting clinical features compatible with renal hypouricaemia. **b**, Disease-associated URAT1 mutants from the patients fail to display urate transport activity. The uptake of urate (300 μM) was measured in oocytes injected with wild-type or mutated cRNA. Data are mean $\pm$ s.e.m. with seven or eight oocytes per group. **c**, Subcellular localization of the wild type and mutants in oocytes. Immunodetection with a specific antibody raised against the C terminus of URAT1 showed that the wild-type, T217M and E298D proteins are expressed at the plasma membrane (arrowheads), whereas fluorescence levels were undetectable in oocytes injected with water or those injected with W258stop cRNA.

associated with the heterogeneity of circulating levels of blood urate among humans and/or with the tendency for hyperuricaemia in patients with gout.

We have characterized a cDNA from the human kidney that encodes a urate transporter. To our knowledge, URAT1 is the first urate transporter to be reported, despite other reports suggesting that galectin 9 (UAT) may function as a urate transporter or channel in cells of numerous tissues²⁸. Galectin 9 has been proposed to have four membrane-spanning domains, but it does not have the typical architecture of a transporter, and it is not clear that urate is translocated via galectin 9 across the cell membrane²⁸. The identification of URAT1 provides insights into renal urate-handling processes and the regulation of circulating urate levels. The prevalence of hyperuricaemia is increasing in the last decade, and it is associated with a risk of developing gout, hypertension, cardiovascular disease and renal failure²⁹. URAT1 displays selective substrate specificity compared with other multispecific OATs^{12–14} and OCTs^{16,17}, making this transporter an attractive target for the development of new drugs against hyperuricaemia. Conventionally, urate is considered an inert end product of purine degradation in humans and higher primates without any physiological value. However, selective evolutionary advantages resulting from the loss of uricase^{7,8} and disadvantages resulting from hyperuricaemia are mutually exclusive. Other studies have demonstrated the biological role of this ubiquitous compound^{1–5}; it is also proposed to be a factor that determines the longevity of primates³⁰. Why humans require such an effective urate reabsorption mechanism in their kidneys should be further elucidated. □

Methods

Cloning of URAT1

We identified various DNA sequences or fragments that were similar to those of the genes for human OATs^{12–14} by searching the publicly available human genome sequence¹⁵ using the BLAST program (<http://www.ncbi.nlm.nih.gov/BLAST>). We identified a sequence that seemed to constitute an exon–intron organization near the OAT4 gene (*SLC22A11*)¹⁴ within a BAC clone, RP11-141J21 (GenBank accession number AC044790), in human chromosome 11q13. The DNA fragments of the predicted gene sequence matched an expressed sequence tag (EST) from the human kidney cDNA library, adding confidence that this search process can identify functional genes. We predicted the initiation codon of this gene (*SLC22A12*) by comparing it with OAT4 cDNA, constructed a gene-specific primer (GSP) upstream of the initiation codon (5'-GCCCGAGTCTGTGAAGCCTA-3'), and succeeded in isolating the full-length cDNA (URAT1) from human kidney poly(A)⁺ RNA by 3'-RACE (Invitrogen). The PCR product was confirmed to be free of errors by comparing its sequence with the human genome sequence.

Tissue distribution

A human multiple-tissue northern blot (MTN) was purchased from Clontech. For northern blot analysis, the full-length URAT1 cDNA was random-primed with [³²P]dCTP. Hybridization was performed at 70 °C for 1 h in ExpressHyb hybridization buffer (Clontech) and washed under a high-stringency condition (0.1 × SSC and 0.1% SDS at 70 °C).

Protein data

For immunohistochemical and western blot analysis, we established a rabbit polyclonal antibody against a keyhole limpet haemocyanin-conjugated synthesized peptide, TLGNSVLKSTQ, corresponding to the 11 amino acids near the carboxy terminus (542–552 of the amino-acid sequence) of URAT1. The human kidney protein (Protein Medley) and human adult normal renal tissue slides were purchased from Clontech and Biochain, respectively. For immunocytochemistry, *Xenopus laevis* oocytes injected with cRNAs were fixed with paraformaldehyde, and incubated with the anti-URAT1 antibody (1:500), followed by Alexa546-labelled goat anti-rabbit immunoglobulin-γ (IgG) (Wako; diluted 1:200). The sections were observed under a confocal laser scanning microscope (Fluoview FV500, Olympus). Alexa546 fluorescence was excited by a green He Ne laser light of wavelength 543 nm.

Expression and transport assay

URAT1 cDNA was inserted into the expression vector pcDNA 3.1(+) (Invitrogen). *In vitro* transcription and injection of capped cRNA (30–50 ng per oocyte) into oocytes were performed as previously described^{12,14}. Oocytes were maintained in Barth's buffer¹² for 2–3 d at 18 °C before use. The ND96 buffer¹² contained (in mM) 96 NaCl, 2 KCl, 1 MgCl₂, 1.8 CaCl₂ and 5 HEPES buffer (pH 7.4). In the 0 Cl[−] bath, Cl[−] was replaced with an equimolar amount of gluconate. The transport assay and the estimation of the kinetic parameters for urate uptake via URAT1 were performed as previously described^{12,14}. The experiments were performed using three batches of oocytes, and results from the representative

experiments are expressed as mean ± s.e.m. Statistical significance was determined by Student's *t*-tests.

Mutation analysis and construction of mutant cDNA

We selected primer sequences for amplification of exons with inclusion of the intron–exon boundaries from intron sequences after establishing the *SLC22A12* structure. Genomic DNA from patients was amplified by PCR, and the products were sequenced using an ABI PRISM 3100 Genetic Analyzer (PE Biosystems). For genotyping the mutation in control individuals, genomic DNA was analysed by fluorophore probe-based techniques to detect single point mutations using real-time fluorescence PCR with a LightCycler (Roche). For the construction of mutants, we used a QuikChange site-directed mutagenesis kit (Stratagene). The mutagenic oligonucleotide primers for generation of W258stop, T217M and E298D mutants were 5'-GGTGTGCGGGACTG(A)ACACTGCTGCAGC-3', 5'-CGTCATGATGAACA(T)GGGCACTCTCCTGATG-3', and 5'-GGGGCTGCAGGA(T)CTGTGGAGGGTGGCTG-3', respectively (sense strands; mutated nucleotides are in parentheses). Proper construction of the mutated cDNA was confirmed by complete sequencing.

Received 3 December; accepted 20 March 2002.

Published online 14 April 2002, DOI 10.1038/nature742.

- Halliwell, B. in *Handbook of Antioxidants* (eds Cadenas, E. & Packer, L.) 243–255 (Dekker, New York, 1996).
- Ames, B. N., Cathcart, R., Schwiers, E. & Hochstein, P. Uric acid provides an antioxidant defence in humans against oxidant- and radical-caused aging and cancer: A hypothesis. *Proc. Natl Acad. Sci. USA* **78**, 6858–6862 (1981).
- Becker, B. F. Towards the physiological function of uric acid. *Free Radical Biol. Med.* **14**, 615–631 (1993).
- Hooper, D. C. *et al.* Uric acid, a natural scavenger of peroxynitrite, in experimental allergic encephalomyelitis and multiple sclerosis. *Proc. Natl Acad. Sci. USA* **95**, 675–680 (1998).
- Keller, J. N. *et al.* Mitochondrial manganese superoxide dismutase prevents neural apoptosis and reduces ischemic brain injury: suppression of peroxynitrite production, lipid peroxidation, and mitochondrial dysfunction. *J. Neurosci.* **18**, 687–697 (1998).
- Abramson, R. G. & Lipkowitz, M. S. in *Basic Principles in Transport, Comparative Physiology* Vol. 3 (ed. Kinne, R. K. H.) 115–153 (Karger, Basel, 1990).
- Wu, X., Lee, C. C., Muzny, D. M. & Caskey, C. T. Urate oxidase: Primary structure and evolutionary implications. *Proc. Natl Acad. Sci. USA* **86**, 9412–9416 (1989).
- Wu, X., Muzny, D. M., Lee, C. C. & Caskey, C. T. Two independent mutational events in the loss of urate oxidase. *J. Mol. Evol.* **34**, 78–84 (1992).
- Maesaka, J. K. & Fishbane, S. Regulation of renal urate excretion: A critical review. *Am. J. Kidney Dis.* **32**, 917–933 (1998).
- Sica, D. A. & Schoolwerth, A. C. in *The Kidney*, 6th edn (ed. Brenner, B. M.) 680–700 (Saunders, Philadelphia, 2000).
- Berlinger, R. W., Hilton, J. G., Yü, T. F. & Kennedy, T. J. Jr The renal mechanism for urate excretion in man. *J. Clin. Invest.* **29**, 396–401 (1950).
- Sekine, T., Watanabe, N., Hosoyamada, M., Kanai, Y. & Endou, H. Expression cloning and characterization of a novel multispecific organic anion transporter. *J. Biol. Chem.* **272**, 18526–18529 (1997).
- Sweet, D. H., Wolff, N. A. & Pritchard, J. B. Expression cloning and characterization of ROAT1. The basolateral organic anion transporter in rat kidney. *J. Biol. Chem.* **272**, 30088–30095 (1997).
- Cha, S. H. *et al.* Molecular cloning and characterization of multispecific organic anion transporter 4 expressed in the placenta. *J. Biol. Chem.* **275**, 4507–4512 (2000).
- International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).
- Gründemann, D., Gorboulev, V., Gambaryan, S., Veyhl, M. & Koepsell, H. Drug excretion mediated by a new prototype of polyspecific transporter. *Nature* **372**, 549–552 (1994).
- Koepsell, H. Organic cation transporters in intestine, kidney, liver, and brain. *Annu. Rev. Physiol.* **60**, 243–246 (1998).
- Roch-Ramel, F., Werner, D. & Guisan, B. Urate transport in brush-border membrane of human kidney. *Am. J. Physiol.* **266**, F797–F805 (1994).
- Guggino, S. E., Martin, G. J. & Aronson, P. S. Specificity and modes of the anion exchanger in dog renal microvillus membranes. *Am. J. Physiol.* **244**, F612–F621 (1983).
- Kahn, A. M., Branham, S. & Weinman, E. J. Mechanism of urate and *p*-aminohippurate transport in rat renal microvillus membrane vesicles. *Am. J. Physiol.* **245**, F151–F158 (1983).
- Emmerson, B. T. The management of gout. *N. Engl. J. Med.* **334**, 445–451 (1996).
- Cullem, J. H., LeVine, M. & Fiore, J. M. Studies on hyperuricemia produced by pyrazinamide. *Am. J. Med.* **23**, 587–595 (1957).
- Roch-Ramel, F., Guisan, B. & Schild, L. Indirect coupling of urate and *p*-aminohippurate transport to sodium in human brush-border membrane vesicles. *Am. J. Physiol.* **270**, F61–F68 (1996).
- Mandel, L. J. Metabolic substrates, cellular energy production, and the regulation of proximal tubular transport. *Annu. Rev. Physiol.* **47**, 85–101 (1985).
- Kikuchi, Y. *et al.* Patients with renal hypouricemia with exercise-induced acute renal failure and chronic renal dysfunction. *Clin. Nephrol.* **53**, 467–472 (2000).
- Igarashi, T., Sekine, T., Sugimura, H., Hayakawa, H. & Arayama, T. Acute renal failure after exercise in a child with renal hypouricemia. *Pediatr. Nephrol.* **7**, 292–293 (1993).
- Halabe, A. & Sperling, O. Uric acid nephrolithiasis. *Miner. Electrol. Metab.* **20**, 424–431 (1994).
- Lipkowitz, M. S. *et al.* Functional reconstitution, membrane targeting, genomic structure, and chromosomal localization of a human urate transporter. *J. Clin. Invest.* **107**, 1103–1115 (2001).
- Johnson, R. J., Kivlighn, S. D., Kim, Y. G., Suga, S. & Fogo, A. B. Reappraisal of the pathogenesis and consequences of hyperuricemia in hypertension, cardiovascular disease, and renal disease. *Am. J. Kidney Dis.* **33**, 225–234 (1999).
- Cutler, R. G. Antioxidants and aging. *Am. J. Clin. Nutr.* **53**, 373S–379S (1991).

Supplementary Information accompanies the paper on Nature's website (<http://nature.com>).

Acknowledgements

We thank the patients for contributing to this study; Y. Terado for discussions on immunohistochemistry, K. Tachampa and J. Y. Kim for help in characterization of URAT1; A. Toki, M. Takahashi and M. Ikeda for technical assistance; and Merck Research Laboratories for providing losartan and EXP-3174. The anti-URAT1 polyclonal antibody was supplied by Trans Genic Inc. (formerly Kumamoto Immunochemical Laboratory). This work was supported in part by grants from the Japanese Ministry of Education, Science, Sports, Culture and Technology, Grants-in-Aid for Scientific Research, and High-Tech Research Center, the Science Research Promotion Fund of the Japan Private School Promotion Foundation.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to H.E. (e-mail: endouh@kyorin-u.ac.jp). The sequences of URAT1 cDNA and protein have been deposited under GenBank/EBML/DBJ accession number AB071863.

Transgenic anopheline mosquitoes impaired in transmission of a malaria parasite

Junitsu Ito^{*†}, Anil Ghosh^{*†}, Luciano A. Moreira^{*}, Ernst A. Wimmer[‡] & Marcelo Jacobs-Lorena^{*}

^{*} Case Western Reserve University, Department of Genetics, 10900 Euclid Avenue, Ohio 44106-4955, USA

[‡] Lehrstuhl für Genetik, Universität Bayreuth, Universitätsstrasse 30, NW1, D-95447 Bayreuth, Germany

[†] These authors contributed equally to this work

Malaria is estimated to cause 0.7 to 2.7 million deaths per year, but the actual figures could be substantially higher owing to under-reporting and difficulties in diagnosis¹. If no new control measures are developed, the malaria death toll is projected to double in the next 20 years¹. Efforts to control the disease are hampered by drug resistance in the *Plasmodium* parasites, insecticide resistance in mosquitoes, and the lack of an effective vaccine. Because mosquitoes are obligatory vectors for malaria transmission, the spread of malaria could be curtailed by rendering them incapable of transmitting parasites. Many of the tools required for the genetic manipulation of mosquito competence for malaria transmission have been developed. Foreign genes can now be introduced into the germ line of both culicine^{2,3} and anopheline⁴ mosquitoes, and these transgenes can be expressed in a tissue-specific manner^{5,6}. Here we report on the use of such tools to generate transgenic mosquitoes that express antiparasitic genes in their midgut epithelium, thus rendering them inefficient vectors for the disease. These findings have significant implications for the development of new strategies for malaria control.

When a mosquito ingests a blood meal from an infected host, *Plasmodium* gametocytes transform into gametes that mate and differentiate into zygotes and then ookinetes (elongated motile zygotes). Ookinetes cross the midgut epithelium and differentiate into oocysts, which after 10–15 days liberate sporozoites into the haemocoel. The development of the parasite in the mosquito is completed when sporozoites cross the salivary gland epithelium⁷. The mechanism by which the parasite crosses the mosquito epithelia is unknown, but is suspected to be receptor mediated. *In vivo* selection from a library of bacteriophages displaying random 12-amino-acid peptides led to the identification of a peptide—PCQRAIFQSCN (termed SM1 for salivary gland- and midgut-binding peptide 1)—that binds specifically to the two epithelia that

are traversed by the parasite: the distal lobes of the salivary glands and the luminal surface of the midgut⁸. Significantly, SM1 strongly inhibited crossing of the two epithelia by the parasites⁸. These results suggest that if SM1 is produced and secreted into the mosquito gut lumen when an infectious blood meal is ingested, then *Plasmodium* development would be blocked.

We searched for a system to drive the expression of genes that inhibit *Plasmodium* development, and found that the carboxypeptidase (CP) promoter and signal sequence has many desirable attributes. The CP promoter is strongly activated by a blood meal, and the CP signal sequence drives secretion of the protein into the midgut lumen, where the initial stages of *Plasmodium* development take place^{6,9}. We constructed a synthetic gene (termed AgCP[SM1]₄) consisting of four SM1 units joined by 4-amino-acid linkers attached to the CP signal sequence and driven by the gut-specific and blood-inducible CP promoter (Fig. 1a). This gene was inserted into a *piggyBac* vector and transformed into the germ line of the mosquito *Anopheles stephensi*⁴. Of 394 embryos injected, 63 (16.0%) larvae hatched, yielding 33 (8.4%) adults. The adults were distributed into 14 families, of which 2 (families A and B) yielded green fluorescent protein (GFP)-positive progeny (Fig. 1b). Progeny from two separate mosquito lines from each family were analysed by Southern blot hybridization (Fig. 1c). The results indicate that each of the four lines originated from a different integration event. Northern blot analysis indicated that the AgCP[SM1]₄ transgene is rapidly and strongly induced by a blood meal in midguts of transgenic mosquitoes with a peak around 3–6 h (Fig. 1d). This pattern is consistent with that previously observed for genes driven by the *Anopheles gambiae* CP promoter^{6,9}. We investigated the synthesis of the AgCP[SM1]₄ protein by the midgut epithelium by immunofluorescence microscopy. The recombinant protein was detected in the midgut epithelium of mosquitoes dissected at 6 h (data not shown) and 24 h (Fig. 2) after a blood meal, but by 36 h the signal had declined to close to basal level (data not shown). Because ookinetes invade the midgut epithelium

Table 1 Inhibition of oocyst formation in transgenic mosquitoes

Experiment	Oocyst prevalence*	Oocyst intensity†	Inhibition (%)‡
1. Control	80 (16/20)	80.8 (0–273)	–
B6	53 (9/17)	25.3 (0–186)	68.7
2. Control	86 (18/21)	70.3 (0–225)	–
B3	37 (7/19)	7.3 (0–40)	89.6
3. Control	94 (17/18)	63.8 (0–365)	–
B3	35 (7/20)	7.2 (0–80)	88.7
4. Control	89 (17/19)	64.9 (0–292)	–
B3, B6	41 (9/22)	3.3 (0–19)	94.9
5. Control	89 (17/19)	132.6 (0–328)	–
B3, B6	54 (14/16)	26.8 (0–105)	79.8
6. Control	89 (17/19)	95.1 (0–290)	–
B3, A3	50 (11/22)	22.1 (0–85)	76.8
7. Control	90 (18/20)	83.4 (0–285)	–
A15	33 (7/21)	9.6 (0–98)	88.5
8. Control	90 (18/20)	129.0 (0–250)	–
A15	45 (10/22)	34.0 (0–134)	73.6
9. Control	86 (19/22)	115.2 (0–292)	–
A15	70 (16/23)	30.3 (0–101)	73.7
Average			
Control	88.1 (17.4/19.8)	93.1 (0–365)	–
Transgenic	46.4 (10.0/21.3)	18.5 (0–186)	81.6

For each experiment, transgenic mosquitoes and sibling control (non-transgenic) mosquitoes from the same rearing were fed simultaneously on the same mouse, which was infected with *P. berghei* ANKA 2.34. A3, A15, B3 and B6 indicate the transgenic lines used in each experiment (Fig. 1c). Where two lines are indicated, a mixture of mosquitoes from those lines was used. All transgenic lines were kept as heterozygotes and mosquitoes were fed on mice with 10–15% parasitaemia and 1–1.5% gametocytaemia. Mosquitoes were kept at 21 °C and the number of oocysts per midgut was counted on day 15 after feeding.

* The per cent mosquitoes that had oocysts in their midgut. This value was derived from the number of oocyst-positive mosquitoes over the total number of mosquitoes examined (shown in parentheses).

† The mean oocyst number per midgut. The range of observed values is indicated in parentheses. In all cases, transgenic mosquito values were significantly different ($P < 0.05$) from those of controls, as analysed by the Mann–Whitney *U*-test.

‡ Reflects the reduction in the mean oocyst number in transgenic mosquitoes relative to control mosquitoes.

around 24 h after a blood meal, it is important that synthesis and secretion of the recombinant peptide precede the time of parasite invasion.

Previous experiments indicated that when an infectious blood meal was fed along with the SM1 peptide, formation of oocysts, but not of ookinetes, was inhibited⁸. To measure the consequences of *AgCP*[SM1]₄ transgene expression on parasite development, we fed control and transgenic mosquitoes on the same infected mouse and measured the numbers of oocysts formed. In nine experiments, inhibition of oocyst formation ranged between 68.7 and 94.9% (average inhibition 81.6%; Table 1). To ascertain that control and transgenic mosquito lines had the same genetic background, the four transgenic lines were backcrossed in each generation to the wild-type mosquito population. We considered the possibility that the observed effects were caused by the fortuitous disruption of an

endogenous mosquito gene on transgene integration or by some other property of the transposon. Two lines of evidence argue against these possibilities. First, equivalent inhibition of oocyst formation was observed with mosquitoes of three independently derived lines (Table 1). Note that for each line, the transgene integrated in a different position in the mosquito genome (Fig. 1c). Second, development of *Plasmodium berghei* in transgenic *A. stephensi* that express GFP from a Minos-based transposon was indistinguishable from development of *P. berghei* in wild-type mosquitoes (F. Catteruccia, personal communication). Thus, the presence of foreign DNA or expression of GFP by themselves do not affect parasite development. Moreover, the SM1 peptide, but not a control (unrelated) peptide, strongly inhibited parasite development and transmission when administered to mosquitoes⁸. These observations suggest that the sequence of the expressed peptide is

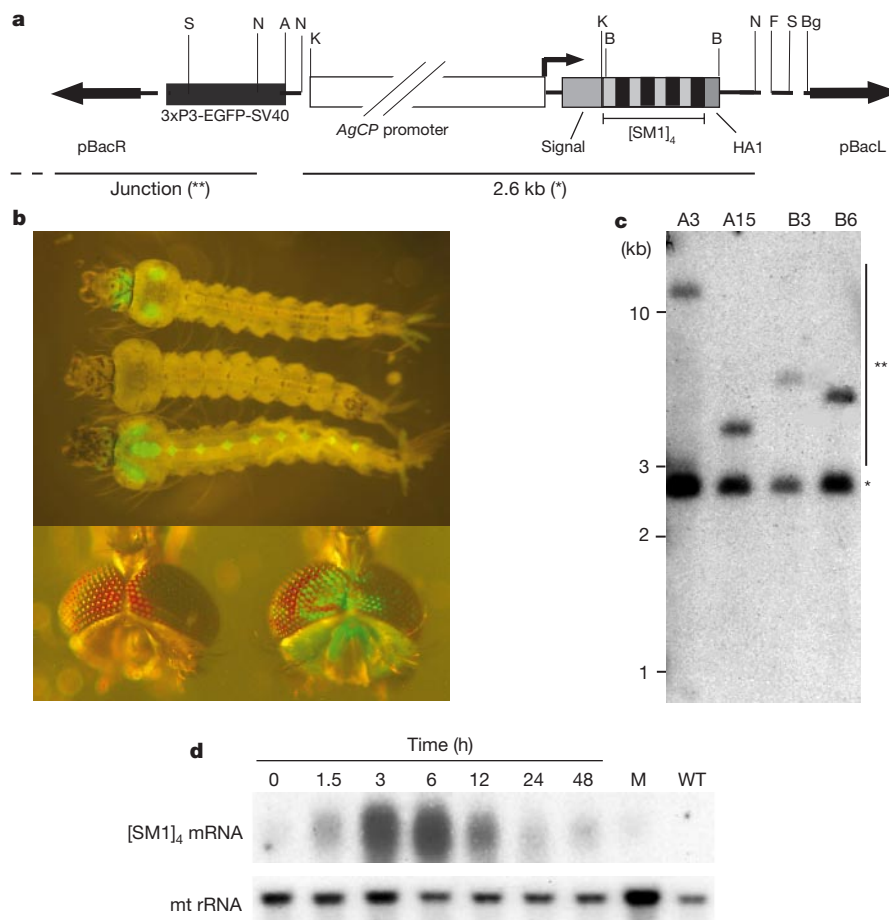


Figure 1 Structure of the *AgCP*[SM1]₄ gene and its expression in transgenic mosquitoes. **a**, Schematic diagram of the *AgCP*[SM1]₄ gene that was transformed into the *A. stephensi* germ line. The construct consists of the *A. gambiae* carboxypeptidase (*AgCP*) promoter (the bent arrow indicates the transcription initiation site), the *AgCP* 5' UTR (line to the right of the promoter), the *AgCP* signal sequence, four units of the SM1 repeat (hatched boxes are the linker amino acids, black boxes are the SM1 peptides), the haemagglutinin epitope (HA1) and the *AgCP* 3' UTR (line to the right of HA1). 3xP3-EGFP-SV40 is the gene that expresses GFP from an eye-specific promoter¹³. The arrows at the end of the construct represent the piggyBac arms. Dashed lines represent flanking plasmid sequences. Restriction sites: S, *Sal*I; N, *Not*I; A, *Asc*I; K, *Kpn*I; B, *Bam*HI; F, *Fse*I; Bg, *Bgl*II. The lines below the construct show the fragments observed in **c**. The size of the junction fragment is variable and depends on the site of integration in the *A. stephensi* genome. **b**, Detection of *AgCP*[SM1]₄ transgenic mosquitoes by transformation marker-mediated fluorescence. Top, a wild-type (non-transgenic) larva (middle) flanked

by transgenic larvae viewed from the dorsal (top) or ventral (bottom) sides. Note green fluorescence of the ventral nerve cord in the latter, which is similar to marker-mediated fluorescence in *Drosophila*¹³. Bottom, the head of a wild-type (left) and a transgenic (right) mosquito. The entire eye expresses GFP but which facets fluoresce depends on the angle of the incident light. **c**, Southern blot analysis of genomic DNA extracted from mosquitoes from two A and two B transgenic lines, digested with *Not*I and *Bgl*II enzymes. The probe was a mixture of [SM1]₄ and 3xP3-EGFP-SV40 sequences (compare with **a**). **d**, Time course of [SM1]₄ messenger RNA accumulation after blood feeding. RNAs were extracted from transgenic female mosquitoes at the times after a blood meal indicated on top of each lane. The RNAs were fractionated by electrophoresis on an agarose gel, blotted onto a nylon membrane and sequentially hybridized first with an [SM1]₄ probe and then with a mitochondrial ribosomal RNA (mt rRNA) probe¹⁴ to verify the amount of RNA analysed in each lane. M, RNA from transgenic male mosquitoes; WT, RNA from wild-type (non-transgenic) female mosquitoes extracted 3 h after a blood meal.

important and that inhibition of *P. berghei* development in the mosquito can be attributed to SM1 expression, not to the transforming vector. SM1 is presumed to bind to a mosquito midgut receptor that is also required for ookinete invasion⁸. It seems that the SM1 tetramer binds to the luminal surface of the midgut (Fig. 2b), inhibiting parasite–epithelium interactions and midgut invasion.

Transgenic mosquitoes were less susceptible to infection (oocyst load) and had fewer sporozoites in their salivary glands than control mosquitoes (Tables 1 and 2). Also, vector competence of transgenic mosquitoes was severely impaired. In two of three experiments, no transmission was detected, and in a third, transmission was reduced by more than twofold (Table 2). In the field, where most mosquitoes carry fewer than five oocysts¹⁰, inhibition of transmission might be very effective. We also note that all experiments were performed with heterozygous mosquitoes that had one copy of the transgene. Inhibition is expected to be even more effective in homozygous mosquitoes that have two copies of the transgene. We were surprised to find that even mosquitoes that had salivary gland sporozoites did not transmit, as indicated by the smaller number of infected mice than infected mosquitoes (Table 2). It is possible that the lightly infected salivary glands had no sporozoites in the duct lumen.

Expression of the SM1 peptide in the mosquito midgut severely reduced vector competence by inhibiting *Plasmodium* development. To our knowledge, this is the first report on the blocking of malaria parasite transmission by a transgenic approach. Preliminary results indicate that the peptide does not alter mosquito fitness (longevity and egg production; unpublished observations). However, many challenges remain to achieve the long-term goal of controlling malaria transmission by genetic modification of the mosquito. A major obstacle will be to devise safe means of spreading

Table 2 Reduction of vector competence in transgenic mosquitoes

Experiment	Sporozoite prevalence*	Sporozoite intensity†	Vector competence‡
1. Control	70 (7/10)	2,320 (0–18,000)	60 (6/10)
A3, B3	13 (1/8)	40 (0–400)	0 (0/8)
2. Control	80 (16/20)	870 (0–4,000)	55 (11/20)
B3, A15	15 (2/13)	62 (0–400)	0 (0/13)
3. Control	80 (8/10)	1,280 (0–3,200)	70 (7/10)
A15	50 (5/10)	240 (0–800)	30 (3/10)

For each experiment, transgenic mosquitoes and sibling control (non-transgenic) mosquitoes were fed on the same mouse, which was infected with *P. berghei*. To measure transmission, single mosquitoes were fed on individual naive mice 25 days after ingesting the infectious blood meal. The salivary gland of each mosquito was dissected immediately after feeding on the mouse, and the number of sporozoites per salivary gland was counted ('sporozoite intensity'). The infection status of each mouse was established by examining a smear of tail vein blood on alternate days. Mice that had no parasites by day 25 were considered not to be infected. A3, A15, B3 and B6 indicate the transgenic lines used in each experiment (Fig. 1c). Where two lines are indicated, a mixture of mosquitoes from those lines was used.

*The per cent mosquitoes that had infected salivary glands. This value was derived from the number of sporozoite-positive mosquitoes over the total number of mosquitoes examined (shown in parentheses).

†The mean sporozoite number per salivary gland. The range of values is indicated in parentheses. This is a minimum estimate because sporozoites from only an aliquot of the salivary gland homogenate were counted. In all cases, infection intensity of transgenic mosquitoes was significantly different ($P < 0.05$) from that of controls, as analysed by the Mann–Whitney *U*-test.

‡The per cent mosquitoes that transmitted the parasite to a naive mouse. The number of infected mice over the total is given in parentheses.

foreign genes across mosquito populations in the field. Another potential obstacle is the genetic diversity and mutability of *Plasmodium*. Because development of the parasite in transgenic mosquitoes is not completely blocked, the possibility exists that 'resistant' variants will be selected. To address this concern, it will be important that mosquitoes be modified with multiple genes, each of which inhibits parasite development by a different mechanism. Work in progress in our laboratory, and in others, is seeking to identify such additional 'effector genes'. Although considerable efforts are needed to respond to these many challenges, the potential payoff is large.

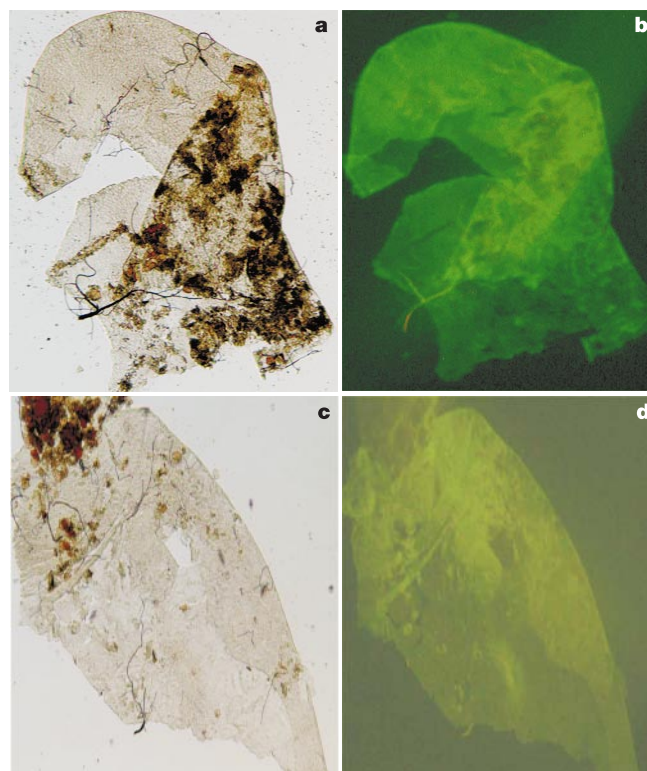


Figure 2 Detection of mosquito-synthesized [SM1]₄ protein. Midguts were dissected 24 h after a blood meal, opened into a sheet, fixed and incubated with an anti-HA1 antibody (Boehringer Mannheim; 1:4,000 dilution), followed by incubation with a fluorescent secondary antibody. **a, b**, Midgut from a female heterozygous for the

AgCP[SM1]₄ gene. **c, d**, Midgut from a wild-type (non-transgenic) female. In each case, a differential interference contrast microscopic image (left) is paired with a fluorescent image (right) of the same midgut.

Genetic manipulation of mosquito vector competence of the type reported here would add a new weapon to the arsenal (drugs, insecticides and perhaps vaccines) for our war against malaria. □

Methods

Transformation vector

For [SM1]₄, a synthetic gene coding for four units of the SM1 peptide (PCQRAIFQSICN) separated by 4-amino-acid (GSPG) linkers was constructed as follows. Two oligonucleotides, SM1⁺ (5'-CCCGTGCCAGCGCGCCATCTTCCAGTCGATCGCAA CGGCTCGCCGGG-3') and SM1⁻ (5'-GCCCGGCGAGCCGTTGCAGATCGACTGGAA GATGGCGCGCTGGCAGCG-3'), were annealed, phosphorylated and self-ligated. The ligation products were fractionated by gel electrophoresis and the 4-repeat unit was excised from the gel to yield [SM1]₄. Two adaptors, 5' and 3', were added to [SM1]₄. The 5' adaptor was obtained by annealing 5'-CGGATCCCCGGG-3' and 5'-GCCCGGGGA TCCGGTAC-3', and the 3' adaptor by annealing 5'-CTACCCCTACGACGTGCCCGAC TACGCCG-3' and 5'-GATCCGCGTAGTCGGGCACGTCGTAGGGGTA-3'. The 3' adaptor codes for the HA1 influenza haemagglutinin epitope.

For 5' Cp, a 1.8-kilobase (kb) *KpnI*-*KpnI* fragment containing the *A. gambiae* CP 1.7-kb promoter, 5' untranslated region (UTR) and signal peptide down to nucleotide +125 (ref. 9) was obtained by PCR with T7 (5'-GTAATACGACTCACTATAGGGC-3') and AgCPKpn (5'-GGTACCCTCGGCGCTTCGACACT-3') primers using the pBluescript AgCP genomic subclone⁹ as a template, followed by digestion with *KpnI*.

For 3' Cp, a 555-base pair (bp) fragment containing the CP 3' region (including the stop codon and 3' UTR; nucleotides +1,337 to +1,880) was obtained by PCR with the primers AgCP3BH (5'-GGATCCTGAAGTCTCTCTACCGG-3') and AgCP3Sc (5'-CCGCGTAAGGCTAGCATTGCCA-3') using the AgCP pBluescript genomic subclone as a template, followed by digestion with *Bam*HI and *Sac*II. The three fragments, 5' Cp, [SM1]₄ with adaptors, and 3' Cp, were combined and sub-cloned into pGEM-T Easy vector (Promega), then digested with *Not*I and inserted into the *Not*I site of pSLfa1180fa (ref. 11). This construct was digested with *Fse*I and *Asc*I, and inserted into the *Fse*I-*Asc*I site of pBac[3xP3-EGFPafm] plasmid¹¹ to yield pBacAgCP[SM1]₄.

Germline transformation

Germline transformation of *A. stephensi* embryos was as previously described⁴ with modifications. Briefly, embryos were treated with 0.2 mM *p*-nitro phenyl *p*'-guanidinobenzoate (Sigma) and microinjected with quartz needles pulled on a P-2000 puller (Sutter). The construct pBacAgCP[SM1]₄ (0.5 mg ml⁻¹) was mixed with the helper phsp-pBac (0.3 mg ml⁻¹)¹². For each founder family, 1–4 adult mosquitoes originating from the injected embryos (G₀) were mated with 5–10 wild-type mosquitoes of the opposite sex. In the next generation (G₁), transgenic mosquitoes were screened by searching for larvae displaying green fluorescence (Fig. 1b). In each generation, mosquitoes were propagated by crossing transgenic males with virgin non-transgenic females from the population that was used to create the transgenic lines. This ensured that the genetic background of all transgenic lines was the same as that of the wild-type control mosquitoes.

Received 28 December 2001; accepted 11 March 2002.

1. Breman, J. G. The ears of the hippopotamus: manifestations, determinants, and estimates of the malaria burden. *Am. J. Trop. Med. Hyg.* **64s**, 1–11 (2001).
2. Jasinskiene, N. *et al.* Stable transformation of the yellow fever mosquito, *Aedes aegypti*, with the Hermes element from the housefly. *Proc. Natl Acad. Sci. USA* **95**, 3743–3747 (1998).
3. Coates, C. J., Jasinskiene, N., Miyashiro, L. & James, A. A. Mariner transposition and transformation of the yellow fever mosquito, *Aedes aegypti*. *Proc. Natl Acad. Sci. USA* **95**, 3748–3751 (1998).
4. Catteruccia, F. *et al.* Stable germline transformation of the malaria mosquito *Anopheles stephensi*. *Nature* **405**, 959–962 (2000).
5. Kokoza, V. *et al.* Engineering blood meal-activated systemic immunity in the yellow fever mosquito, *Aedes aegypti*. *Proc. Natl Acad. Sci. USA* **97**, 9144–9149 (2000).
6. Moreira, L. A. *et al.* Robust gut-specific gene expression in transgenic *Aedes aegypti* mosquitoes. *Proc. Natl Acad. Sci. USA* **97**, 10895–10898 (2000).
7. Ghosh, A., Edwards, M. J. & Jacobs-Lorena, M. The journey of malaria in the mosquito: hopes for the new century. *Parasitol. Today* **16**, 196–201 (2000).
8. Ghosh, A., Ribolla, P. E. M. & Jacobs-Lorena, M. Targeting *Plasmodium* ligands on mosquito salivary glands and midgut with a phage display peptide library. *Proc. Natl Acad. Sci. USA* **98**, 13278–13281 (2001).
9. Edwards, M. J., Lemos, F. J., Donnelly-Doman, M. & Jacobs-Lorena, M. Rapid induction by a blood meal of a carboxypeptidase gene in the gut of the mosquito *Anopheles gambiae*. *Insect Biochem. Mol. Biol.* **27**, 1063–1072 (1997).
10. Pringle, G. A quantitative study of naturally-acquired malaria infections in *Anopheles gambiae* and *Anopheles funestus* in a highly malarious area of East Africa. *Trans. R. Soc. Trop. Med. Hyg.* **60**, 626–632 (1966).
11. Horn, C. & Wimmer, E. A. A versatile vector set for animal transgenesis. *Dev. Genes Evol.* **210**, 630–637 (2000).
12. Handler, A. M. & Harrell, R. A. II Germline transformation of *Drosophila melanogaster* with the piggyBac transposon vector. *Insect Mol. Biol.* **8**, 449–457 (1999).
13. Horn, C., Jaunich, B. & Wimmer, E. A. Highly sensitive, fluorescent transformation marker for *Drosophila* transgenesis. *Dev. Genes Evol.* **210**, 623–629 (2000).
14. Lemos, F. J. A., Cornel, A. J. & Jacobs-Lorena, M. Trypsin and aminopeptidase gene expression is affected by age and food composition in *Anopheles gambiae*. *Insect Biochem. Mol. Biol.* **26**, 651–658 (1996).

Acknowledgements

We thank J. Snyder and G. Hundemer for help, and members of the laboratory for comments. This investigation received financial support from the UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases (TDR) and from the National Institutes of Health. E.A.W. acknowledges support by the Robert Bosch Foundation.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to M.J.-L. (e-mail: mxj3@po.cwru.edu).

HDAC6 is a microtubule-associated deacetylase

Charlotte Hubbert*, Amaris Guardiola*, Rong Shao*†, Yoshiharu Kawaguchi*‡, Akihiro Ito*, Andrew Nixon*, Minoru Yoshida†, Xiao-Fan Wang* & Tso-Pang Yao*

* Department of Pharmacology and Cancer Biology, Duke University, Durham, North Carolina 27710, USA

† Department of Biotechnology, The University of Tokyo, Bunkyo-ku, Tokyo 113-8657, Japan; CREST Research Project, Japan Science and Technology Corporation, Saitama 332-0012, Japan

‡ These authors contributed equally to this work

Reversible acetylation of α -tubulin has been implicated in regulating microtubule stability and function¹. The distribution of acetylated α -tubulin is tightly controlled and stereotypic. Acetylated α -tubulin is most abundant in stable microtubules but is absent from dynamic cellular structures such as neuronal growth cones and the leading edges of fibroblasts^{1,2}. However, the enzymes responsible for regulating tubulin acetylation and deacetylation are not known. Here we report that a member of the histone deacetylase family, HDAC6, functions as a tubulin deacetylase. HDAC6 is localized exclusively in the cytoplasm, where it associates with microtubules and localizes with the microtubule motor complex containing p150^{glued} (ref. 3). *In vivo*, the overexpression of HDAC6 leads to a global deacetylation of α -tubulin, whereas a decrease in HDAC6 increases α -tubulin acetylation. *In vitro*, purified HDAC6 potentially deacetylates α -tubulin in assembled microtubules. Furthermore, overexpression of HDAC6 promotes chemotactic cell movement, supporting the idea that HDAC6-mediated deacetylation regulates microtubule-dependent cell motility. Our results show that HDAC6 is the tubulin deacetylase, and provide evidence that reversible acetylation regulates important biological processes beyond histone metabolism and gene transcription.

Extensive studies of histone acetylation, a process controlled by histone acetyltransferases (HATs) and histone deacetylases (HDACs), have firmly established a role for reversible acetylation in transcriptional regulation and histone metabolism⁴. At least 11 proteins predicted to be members of the HDAC family have been identified on the basis of homology within the catalytic domain^{5,6}. The sequences outside the catalytic domain are highly divergent, indicating that these enzymes might have different biological functions and a broader substrate repertoire beyond histones. Indeed, recent studies reveal that many non-histone nuclear transcription factors, such as p53, E2Fs and myoD, are regulated by acetylation^{7–9}. Furthermore, there are also cytoplasmic proteins that are subject to modification by acetylation (reviewed in ref. 10). The most notable of these is α -tubulin.

To examine whether acetylation might regulate cellular functions unrelated to histone metabolism and transcriptional regulation, we focused on HDAC6. HDAC6 is localized exclusively in the cytoplasm^{11,12} (Fig. 1) and therefore seemed likely to have unique functions. Immunostaining reveals that endogenous HDAC6 is localized to punctate structures concentrated perinuclearly, as well as the leading edge of the cell (Fig. 1a). This unique cytoplasmic distribution is reminiscent of the patterns generated by p150^{glued} (ref. 3), a protein found in the dynein–dynactin microtubule motor complex. Double immunostaining with antibodies against HDAC6 and p150^{glued} shows significant co-localization of these proteins around the nucleus and at the leading edge (Fig. 1b–d). These results indicate that HDAC6 is a microtubule-associated deacetylase. Supporting this idea, nocodazole treatment, which collapses the microtubule network, leads to the redistribution of HDAC6 around the nucleus, where it completely co-localizes with p150^{glued} (Fig. 1e–g) and tubulin (Fig. 1k–m). The same treatment has no effect on the actin cytoskeleton (Fig. 1n, o). Although HDAC6 in untreated cycling cells shows only marginal co-localization with microtubules (Fig. 1h–j), co-localization of HDAC6 and the microtubule network is observed in quiescent cells (Fig. 1p–r). Thus, HDAC6 is a cytoplasmic deacetylase that associates with microtubules.

Because acetylation of α -tubulin is an important post-translational modification of the microtubule network, we examined whether HDAC6 functions as a tubulin deacetylase. NIH-3T3 cells, which normally express low levels of endogenous HDAC6, were generated to overexpress wild-type or catalytically inactive HDAC6 stably, to determine whether HDAC6 modulates tubulin

acetylation (Fig. 2a). As an additional control, a stable line expressing the class I histone deacetylase HDAC1, was also evaluated. The effect of HDAC6 overexpression on the acetylation status of α -tubulin *in vivo* was then analysed. To evaluate the abundance of acetylated tubulin, we used an antibody (6-11B-1) that specifically recognizes tubulin only when it is acetylated. The specificity of this antibody was determined by its reactivity with α -tubulin that was naturally or chemically acetylated on Lys 40, but not with α -tubulin that was unmodified¹³. Moderate levels of acetylated tubulin were observed in control NIH-3T3 cells lines (Fig. 2b, lane 1). In the cells that overexpress wild-type HDAC6 (Fig. 2b, lane 3) there is a clear reduction in acetylated tubulin levels, which is not observed in the cells overexpressing the catalytically inactive HDAC6 or HDAC1 (Fig. 2b, lanes 2 and 4). This result shows that overexpression of HDAC6 specifically reduces the acetylation levels of α -tubulin. To confirm this result, we tested HDAC6 activity in the presence of the potent cancer therapeutic agent Taxol, which promotes microtubule stability and induces high levels of α -tubulin acetylation¹ (Fig. 2c). Overexpression of HDAC6 markedly reduces the levels of tubulin acetylation induced by Taxol (Fig. 2c, lane 6). These results further support the idea that HDAC6 functions as a tubulin deacetylase *in vivo*.

Unlike other family members, HDAC6 is uniquely resistant to the potent HDAC inhibitors trapoxin-B (TPXb)¹⁴ and sodium butyrate as a result of its dual-catalytic-domain structure⁶. Like all other HDACs, however, HDAC6 is inhibited by trichostatin A (TSA). We used these drugs to test further the specificity of the overexpressed and endogenous activity of HDAC6 towards acetylated α -tubulin.

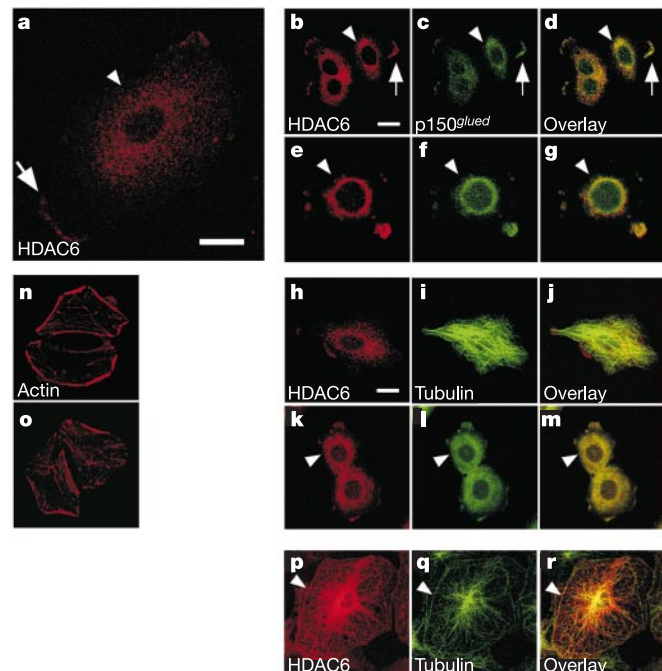


Figure 1 Co-localization of HDAC6 with p150^{glued} and the microtubule network. **a–g**, Immunolocalization of endogenous HDAC6 and p150^{glued} in A549 cells (**a–d**) and in nocodazole-treated cells (**e–g**) as indicated. **h–m**, Immunolocalization of HDAC6 and α -tubulin in the absence (**h–j**) or presence (**k–m**) of nocodazole. **n, o**, Actin network (phalloidin stain, red) is unaffected in the absence (**n**) or presence (**o**) of nocodazole. **p–r**, Localization of HDAC6 and α -tubulin in serum-starved NIH-3T3 cells stably overexpressing HDAC6. Co-localization is documented by superimposing (overlay) individual images generated from three-dimensional projections of confocal Z-sections. Scale bar, 10 μ m. Arrowheads indicate concentration of staining in the perinuclear region or the collapsed microtubule network; arrows indicate the leading edge.

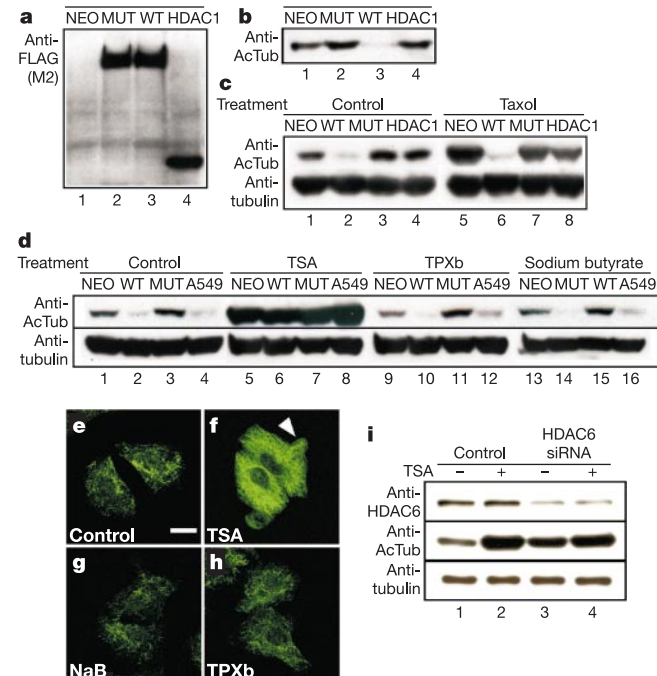


Figure 2 HDAC6 is a tubulin deacetylase *in vivo*. **a**, Levels of wild-type HDAC6 (WT), catalytically inactive HDAC6 (MUT), HDAC1 and control (NEO) in stable NIH-3T3 cells by immunoblotting with anti-Flag (M2) antibody. **b**, Acetylated α -tubulin levels (AcTub) in various NIH-3T3 lines as indicated. **c**, Acetylated and total α -tubulin in untreated NIH-3T3 lines (lanes 1–4) or NIH-3T3 lines treated with paclitaxel (lanes 5–8). **d**, Acetylated and total α -tubulin in NIH-3T3 lines and A549 cells treated with TSA, TPXb or sodium butyrate as indicated. **e–h**, Acetylated α -tubulin immunostaining in A549 cells treated with TSA, TPXb or sodium butyrate as indicated. **i**, A549 cells were mock-transfected (control) or transfected with HDAC6 siRNA (nucleotides 211–231) then treated with TSA as indicated. Protein levels were determined by immunoblotting with specific antibodies.

As shown in Fig. 2d, treatment with TSA, but not TPXb or sodium butyrate, completely reversed the effect of HDAC6 on tubulin acetylation levels. These findings support the conclusion that the observed reduction in tubulin acetylation is mediated by HDAC6. Interestingly, the levels of acetylated tubulin in all of the stable cell lines exceed basal levels with TSA treatment (Fig. 2d, lanes 5–8), indicating that TSA also inhibits endogenous tubulin deacetylase activity. Consistent with this idea, TSA treatment increases the acetylated tubulin levels in A549 cells, whereas treatment with sodium butyrate or TPXb has no effect (Fig. 2d). To corroborate these biochemical results, we directly visualized changes in acetylated tubulin in response to various drug treatments by immunostaining. A marked increase in the acetylation of the microtubule network is clearly evident in cells treated with TSA but not in cells treated with sodium butyrate or TPXb (Fig. 2e–h). Taken together, these results confirm that HDAC6 shares unique pharmacological properties with the endogenous tubulin deacetylase. To demonstrate further that HDAC6 is indeed the endogenous tubulin deacetylase, we used short interfering RNA (siRNA) to inhibit its expression. As shown in Fig. 2i, a decrease in HDAC6 leads to a marked increase (~5-fold) of acetylated α -tubulin. In cells with reduced expression of HDAC6, treatment with TSA results in only a marginal increase in α -tubulin acetylation (compare lanes 3 and 4 with 1 and 2). We conclude from these experiments that TSA-induced tubulin acetylation is a result of the inhibition of HDAC6. These sets of experiments demonstrate that HDAC6 is the endogenous tubulin deacetylase.

We next determined whether HDAC6 could directly deacetylate acetylated tubulin *in vitro*. Tubulin deacetylation was evaluated by incubating recombinant HDAC6 protein with tubulin prepared from brain tissue, which is heavily acetylated. Reactions were analysed for levels of acetylated α -tubulin. To our surprise, HDAC6 has no deacetylase activity towards free α/β -tubulin dimers (Fig. 3a). As acetylation is generally observed on more stable microtubules, we reasoned that polymerized microtubules might be the preferred substrate for HDAC6. To test this hypothesis, we

assembled tubulin preparations enriched with microtubule-associated proteins (MAPs) into microtubules and then incubated with recombinant HDAC6. Under these conditions, HDAC6 incubation leads to efficient tubulin deacetylation. Neither the catalytically inactive HDAC6 nor HDAC1 has activity against acetylated microtubules (Fig. 3b, c). Consistent with the analysis *in vivo* is the observation that the tubulin deacetylase activity of HDAC6 was completely inhibited by TSA but was insensitive to TPXb (Fig. 3c, lanes 4–6). These findings show that HDAC6 is able to deacetylate a polymerized microtubule substrate *in vitro* and exhibits the same pharmacological specificity as that demonstrated *in vivo*.

Acetylated microtubules are widely observed to represent a more stable microtubule population. Consistent with this observation is the fact that acetylated microtubules are absent from the fibroblast leading edge¹, a highly dynamic structure devoid of stable microtubules and involved in cell motility¹⁵, whereas HDAC6 is enriched in this region (Fig. 1a, b). These observations raise the possibility that HDAC6 might regulate cell motility by controlling microtubule acetylation at the leading edge. We therefore determined whether overexpression of HDAC6 changes the chemotactic movement of NIH-3T3 cells. As shown in Fig. 4, cells overexpressing wild-type HDAC6 move at least 3.5-fold faster than control NIH-3T3 cells in response to serum (Fig. 4b). Cells overexpressing catalytic-inactive HDAC6 show motility similar to that of control cells, showing that the deacetylase activity of HDAC6 is required for an increase in chemotactic cell motility.

Our evidence reveals that HDAC6 localizes to the perinuclear and leading-edge subcellular regions and is associated, in part, with microtubules. The co-localization with the p150^{glued}-containing motor complex indicates that HDAC6 might control microtubule motor-based cargo transport or sorting by modulating the acetylation of specific proteins. Recent studies show that microtubule stability affects actin-dependent membrane ruffling and cell motility by activating the small GTPase Rac-1 (ref. 16) or Rho (ref. 17). We propose that HDAC6 might function at the interface of the

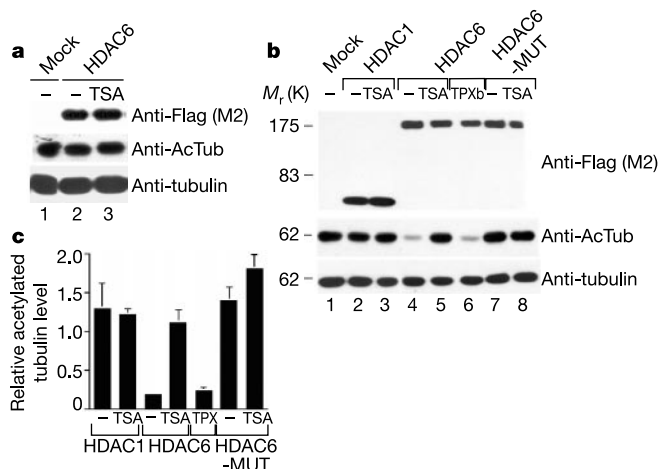


Figure 3 Recombinant HDAC6 exhibits tubulin deacetylase activity towards assembled microtubules *in vitro*. **a**, Acetylated α/β -tubulin dimers were incubated alone (mock), with HDAC6 or with HDAC6 plus TSA as indicated. Acetylated (AcTub) and total α -tubulin levels were determined by immunoblotting. **b**, MAP-stabilized microtubules were incubated alone or with HDAC1, HDAC6 or catalytically inactive HDAC6 (MUT) in the presence or absence of TSA or TPXb as indicated. Top, middle and bottom panels indicate HDAC expression levels, acetylated and total α -tubulin levels, respectively. **c**, Relative acetylated α -tubulin levels were quantified by normalizing acetylated α -tubulin to total α -tubulin levels. Results are an average of three independent experiments performed as in **b**.

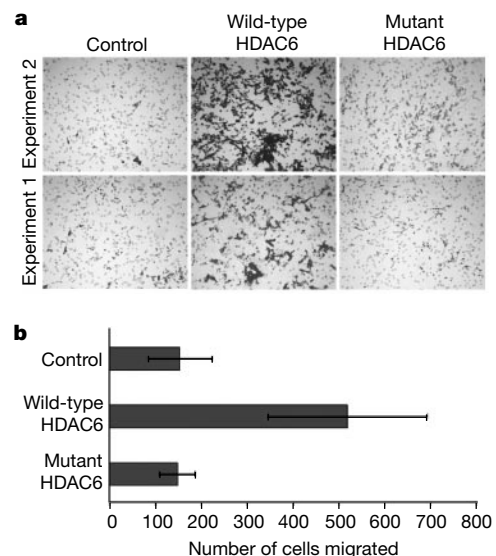


Figure 4 Overexpression of HDAC6 increases the chemotactic motility of NIH-3T3 cells. **a**, NIH-3T3 cells overexpressing control (*neo*-cassette alone), wild-type HDAC6 or catalytically inactive HDAC6 assayed for chemotactic motility in a Transwell assay in response to serum. Cells that traversed the membrane were visualized by a differential haematology stain. Representative fields are shown from two experiments. **b**, Number of cells in the Transwell assay that exhibited chemotactic movement. Results are an average of three independent experiments.

microtubule and actin networks by regulating microtubule acetylation and stability to control cell motility. Furthermore, HDAC6 contains a unique zinc-finger motif that is probably involved in the regulation of ubiquitination¹⁸ (J. Kovacs, J. T. Wu and T.P.Y., unpublished observation), a modification known to be important for protein sorting and transport¹⁹. As ubiquitination and acetylation both target lysine residues, it is probable that HDAC6 could coordinate these two important modifications. If this is true, HDAC6 could have activities in addition to modulating tubulin acetylation. Our study has identified a function for reversible acetylation in cell motility and supports the idea that this post-translational modification has a much broader function in cell signalling and homeostasis. □

Methods

Cell culture, stable lines and drug treatment

To establish stable NIH-3T3 cell lines, retroviruses expressing HDAC6, catalytically inactive HDAC6, HDAC1 or the neomycin-resistance cassette alone were packaged in phoenix cells. Viral supernatants were used to infect NIH-3T3 cells followed by G418 (600 µg ml⁻¹) selection to obtain pools of stable cell lines. All inhibitors except TPXb were obtained from Sigma and were used at the following concentrations for the durations shown: sodium butyrate (5 mM, 8.5 h), TPXb (100 nM, 8 h), TSA (5 µM, 4 h) and paclitaxel (1 µM, 15 h); in Fig. 2, immunostaining experiments with A549 used trichostatin A and TPXb for 1 h and 4 h, respectively. Nocodazole collapse experiments were performed in the presence of 5 µg ml⁻¹ nocodazole for 3 h.

Plasmids and antibodies

Flag-tagged pBJ5-HDAC6 expression plasmids were provided by S. Schreiber (Harvard University). Wild-type and catalytically inactive HDAC6 and HDAC1 retroviral plasmids were made in a CR-Neo retroviral vector. Anti-HDAC6 antibodies were generated by injecting rabbits with recombinant human HDAC6 (amino acids 1044–1273) followed by affinity purification. Antisera from two different rabbits (184 and 185) show identical results in western blotting and in immunohistochemistry.

Immunostaining and fluorescence microscopy

Immunostaining was performed essentially as described previously²⁰ except for experiments visualizing α-tubulin and acetylated α-tubulin, in which cells were fixed in methanol at -20 °C as described previously¹. Antibodies were used at the following dilutions: anti-HDAC6 (1:100), anti-p150^{glued} (Transduction Laboratories; 1:50), anti-α-tubulin (Sigma; 1:100) and anti-acetylated α-tubulin (Sigma; 1:50) as indicated. Actin was stained with rhodamine-phalloidin (Molecular Probes; dilution 1:200) for 10 min. All antibodies (dilutions as given above) were incubated overnight at 4 °C. Laser-scanning confocal microscopy was performed on a Zeiss LSM410 with krypton-argon and helium-neon lasers; three-dimensional projections were generated with the accompanying software.

Immunoblotting

Cells were lysed in buffer (20 mM Tris-HCl at pH 7.6, 170 mM NaCl, 1 mM EDTA, 0.5% Nonidet P40, 1 mM dithiothreitol) supplemented with 5 µM TSA and protease inhibitors. Whole cell lysates (50 µg) were analysed in a western blotting experiment with anti-Flag (M2) antibody (Sigma; dilution 1:1,000), anti-α-tubulin antibody (1:1,000), anti-acetylated α-tubulin antibody (1:1,000) or anti-HDAC6 antibody (1:1,000).

Migration assay

Transwell migration assay was performed as described previously²¹. In brief, stable NIH-3T3 lines were serum starved overnight and plated at 150,000 cells per upper chamber, and 10% calf serum was placed in the lower chamber. Membrane inserts (8.0 µm; Falcon) were coated with 0.1 mg ml⁻¹ collagen IV (Benton-Dickson). The assay was routinely run for 4 h at 37 °C. At completion, membranes were fixed in 10% formalin and stained with Harris Hematoxylin Stain (Sigma); cells were then removed from the upper side of the membrane with a cotton swab before being mounted in Cytoseal 60 (Stephens Scientific). Five random fields were counted per membrane under low power and photographed with a cooled charge-coupled-device camera (Princeton Instruments) interfaced with Metamorph software (Universal Imaging Corp.).

Tubulin deacetylase (TDAC) assay

Flag-tagged HDAC-expressing stable NIH-3T3 lines were lysed in low-stringency lysis buffer (50 mM Tris-HCl at pH 7.6, 120 mM NaCl, 0.5 mM EDTA, 0.5% Nonidet P40, 0.5 mM phenylmethylsulphonyl fluoride, 2 µg ml⁻¹ aprotinin, 1 µg ml⁻¹ leupeptin). Lysates were precleared and immunoprecipitated with anti-Flag (M2) antibody. TDAC assays were performed in a 200-µl volume containing 10 mM Tris-HCl at pH 8, 10 mM NaCl and 50 µg MAP-stabilized microtubules (polymerized from a MAP-rich tubulin fraction) or 35 µg α/β-tubulin dimers. Reactions that included HDAC inhibitors were pretreated with 400 nM TSA or 100 nM TPXb. Reactions were incubated at 37 °C for 2 h then put on ice for 15 min. Reactions were centrifuged to separate pellet-protein-A beads

from the supernatant. The beads were analysed by western blotting with an anti-Flag (M2) antibody and the supernatant was analysed with anti-acetylated α-tubulin and anti-α-tubulin antibodies.

siRNA assay

Two 21-nucleotide double-stranded siRNAs were generated against the amino terminus of HDAC6 at nucleotides 211–231 and 217–237 (Dharmacon Research). A549 cells were transfected with the siRNA at a final concentration of 10 nM per sample by using the TransIT-TKO Transfection Reagent (Mirus Corporation) and in accordance with the manufacturer's instructions. In brief, cells were transfected three times over 5 d and collected 48 h after the final transfection. For TSA-treated samples, cells were treated with 1 µM TSA for 2 h before being collected. Whole cell lysates (35 µg) were analysed in western blotting experiments with anti-HDAC6, anti-acetylated α-tubulin and anti-α-tubulin antibodies. Mock transfection with transfecting reagent alone served as control. Both sets of siRNA yielded the same results, only one of which is shown in Fig. 2i.

Received 28 February; accepted 27 March 2002.

- Piperno, G., LeDizet, M. & Chang, X. J. Microtubules containing acetylated α-tubulin in mammalian cells in culture. *J. Cell Biol.* **104**, 289–302 (1987).
- Robson, S. J. & Burgoyne, R. D. Differential localisation of tyrosinated, detyrosinated, and acetylated α-tubulins in neurites and growth cones of dorsal root ganglion neurons. *Cell Motil. Cytoskel.* **12**, 273–282 (1989).
- Smith, D. S. *et al.* Regulation of cytoplasmic dynein behaviour and microtubule organization by mammalian Lis1. *Nature Cell Biol.* **2**, 767–775 (2000).
- Wade, P. A., Pruss, D. & Wolffe, A. P. Histone acetylation: chromatin in action. *Trends Biochem. Sci.* **22**, 128–132 (1997).
- Grozinger, C. M., Hassig, C. A. & Schreiber, S. L. Three proteins define a class of human histone deacetylases related to yeast Hda1p. *Proc. Natl Acad. Sci. USA* **96**, 4868–4873 (1999).
- Guardiola, A. R. & Yao, T. P. Molecular cloning and characterization of a novel histone deacetylase HDAC10. *J. Biol. Chem.* **277**, 3350–3356 (2001).
- Luo, J., Su, F., Chen, D., Shiloh, A. & Gu, W. Deacetylation of p53 modulates its effect on cell growth and apoptosis. *Nature* **408**, 377–381 (2000).
- Ito, A. *et al.* p300/CBP-mediated p53 acetylation is commonly induced by p53-activating agents and inhibited by MDM2. *EMBO J.* **20**, 1331–1340 (2001).
- Sartorelli, V. *et al.* Acetylation of MyoD directed by PCAF is necessary for the execution of the muscle program. *Mol. Cell* **4**, 725–734 (1999).
- Kouzarides, T. Acetylation: A regulatory modification to rival phosphorylation? *EMBO J.* **19**, 1176–1179 (2000).
- Verdel, A. *et al.* Active maintenance of mHDA2/mHDAC6 histone-deacetylase in the cytoplasm. *Curr. Biol.* **10**, 747–749 (2000).
- Barlow, A. L. *et al.* dSIR2 and dHDAC6: two novel, inhibitor-resistant deacetylases in *Drosophila melanogaster*. *Exp. Cell Res.* **265**, 90–103 (2001).
- Piperno, G. & Fuller, M. T. Monoclonal antibodies specific for an acetylated form of α-tubulin recognize the antigen in cilia and flagella from a variety of organisms. *J. Cell Biol.* **101**, 2085–2094 (1985).
- Furumai, R. *et al.* Potent histone deacetylase inhibitors built from trichostatin A and cyclic tetrapeptide antibiotics including trapoxin. *Proc. Natl Acad. Sci. USA* **98**, 87–92 (2001).
- Wadsworth, P. Regional regulation of microtubule dynamics in polarized, motile cells. *Cell Motil. Cytoskel.* **42**, 48–59 (1999).
- Waterman-Storer, C. M. & Salmon, E. Positive feedback interactions between microtubule and actin dynamics during cell motility. *Curr. Opin. Cell Biol.* **11**, 61–67 (1999).
- Kaverina, I., Krylyshkina, O. & Small, J. V. Microtubule targeting of substrate contacts promotes their relaxation and dissociation. *J. Cell Biol.* **146**, 1033–1044 (1999).
- Seigneurin-Berny, D. *et al.* Identification of components of the murine histone deacetylase 6 complex: Link between acetylation and ubiquitination signalling pathways. *Mol. Cell Biol.* **21**, 8035–8044 (2001).
- Hicke, L. A new ticket for entry into budding vesicles—ubiquitin. *Cell* **106**, 527–530 (2001).
- Eckner, R., Yao, T. P., Oldread, E. & Livingston, D. M. Interaction and functional collaboration of p300/CBP and bHLH proteins in muscle and B-cell differentiation. *Genes Dev.* **10**, 2478–2490 (1996).
- Malinda, K. M., Ponce, L., Kleinman, H. K., Shackleton, L. M. & Millis, A. J. Gp38k, a protein synthesized by vascular smooth muscle cells, stimulates directional migration of human umbilical vein endothelial cells. *Exp. Cell Res.* **250**, 168–173 (1999).

Acknowledgements

We thank A. R. Means, A. M. Pendergast and A. M. J. Van Dongen, as well as the members of the Yao laboratory, for their insightful discussions and critical reading of the manuscript; A. M. Pendergast for pointing out the similarity between HDAC6 and p150^{glued}; and J. Gross and H. Solari for technical assistance with the confocal microscopy. X.F.W. was supported by an NIH grant, as was A.N. (to P. J. Casey). This work is supported by funding from the Damon Runyon–Walter Winchell Cancer Foundation and the Department of Defense to T.-P.Y.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to T.-P.Y. (e-mail: yao00001@mc.duke.edu).

Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities

A. C. R. da Silva¹, J. A. Ferro², F. C. Reinach^{1,20}, C. S. Farah¹, L. R. Furlan³, R. B. Quaggio¹, C. B. Monteiro-Vitorello⁴, M. A. Van Sluys⁵, N. F. Almeida⁶, L. M. C. Alves², A. M. do Amaral⁷, M. C. Bertolini⁸, L. E. A. Camargo⁴, G. Camarotte⁴, F. Cannavan⁹, J. Cardozo¹⁰, F. Chambergo¹, L. P. Ciapina², R. M. B. Cicarelli¹⁰, L. L. Coutinho⁴, J. R. Cursino-Santos¹¹, H. El-Dorry¹, J. B. Faria¹², A. J. S. Ferreira¹, R. C. C. Ferreira¹², M. I. T. Ferro², E. F. Formighieri⁹, M. C. Franco⁹, C. C. Greggio², A. Gruber¹³, A. M. Katsuyama¹³, L. T. Kishi², R. P. Leite¹⁴, E. G. M. Lemos², M. V. F. Lemos¹⁵, E. C. Locali⁷, M. A. Machado⁷, A. M. B. N. Madeira¹³, N. M. F. Martinez-Rossi¹¹, E. C. Martins¹⁰, J. Meidanis¹⁶, C. F. M. Menck¹², C. Y. Miyaki¹⁷, D. H. Moon⁹, L. M. Moreira¹, M. T. M. Novo⁸, V. K. Okura¹⁶, M. C. Oliveira⁵, V. R. Oliveira¹², H. A. Pereira², A. Rossi¹⁸, J. A. D. Sena¹⁵, C. Silva¹³, R. F. de Souza¹, L. A. F. Spinola¹³, M. A. Takita⁷, R. E. Tamura¹³, E. C. Teixeira⁸, R. I. D. Tezza², M. Trindade dos Santos¹, D. Truffi⁴, S. M. Tsal⁹, F. F. White^{2,20}, J. C. Setubal¹⁶ & J. P. Kitajima^{19,20}

¹Departamento de Bioquímica, Instituto de Química, Av. Prof. Lineu Prestes 748;

⁵Departamento de Botânica, Instituto de Biociências, Rua do Matão 277;

¹²Departamento Microbiologia, Instituto de Ciências Biomédicas, Av. Prof. Lineu Prestes 1374; ¹³Departamento de Patologia, Faculdade de Medicina Veterinária e Zootecnia, Av. Prof. Dr. Orlando Marques de Paiva 87; and ¹⁷Departamento de Biologia, Instituto de Biociências, Rua do Matão 277, Universidade de São Paulo, 05508-900, São Paulo, SP, Brazil

²⁵Departamento de Tecnologia, and ¹⁵Departamento de Biologia Aplicada à Agropecuária, Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista, Via de Acesso Prof. Paulo Donato Castellane s/n, 14884-900, Jaboticabal, SP, Brazil

³Departamento de Melhoramento e Nutrição Animal, Faculdade de Medicina Veterinária e Zootecnia, Universidade Estadual Paulista, Fazenda Lageado, Caixa Postal 560, 18600-000, Botucatu, SP, Brazil

⁴Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Av. Pádua Dias 11, 13418-900, Piracicaba, SP, Brazil

⁶Departamento de Computação e Estatística, Universidade Federal de Mato Grosso do Sul, Cx Postal 549, 79070-900, Campo Grande, MS, Brazil

⁷Centro de Citricultura Sylvio Moreira—IAC, Rod. Anhanguera, km 158, Cx. Postal 04, 13.490-970, Cordeirópolis, SP, Brazil

⁸Departamento de Bioquímica e Tecnologia Química, Instituto de Química, Universidade Estadual Paulista, Cx. Postal 355, 14801-970, Araraquara, SP, Brazil

⁹Centro de Energia Nuclear na Agricultura, Av. Centenário 303, 13400-970, Piracicaba, SP, Brazil

¹⁰Departamento de Ciências Biológicas, Faculdade Ciências Farmacêuticas, Universidade Estadual Paulista, Rod. Araraquara-Jaú km 01, 14.801-902, Araraquara, SP, Brazil

¹¹Departamento de Genética, and ¹⁸Departamento de Bioquímica e Imunologia, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Av. Bandeirantes 3900, 14049-900, Ribeirão Preto, SP, Brazil

¹⁴Instituto de Tecnologia do Paraná, IAPAR, Cx. Postal 481, 86001-970, Londrina, PR, Brazil

¹⁶Instituto de Computação, and ¹⁹Centro de Biologia Molecular e Engenharia Genética/Instituto de Computação, Universidade Estadual de Campinas, Cx Postal 6176, 13084-971, Campinas, SP, Brazil

²⁰Present addresses: Aleryx Applied Genomics, Rua James Clark Maxwell 320, Techno Park, 13067-850, Campinas, SP, Brazil (F.C.R. and J.P.K.); Department of Plant Pathology, Kansas State University, 4024 Throckmorton Hall, Manhattan, Kansas 66506, USA (F.F.W.)

The genus *Xanthomonas* is a diverse and economically important group of bacterial phytopathogens, belonging to the γ -subdivision of the Proteobacteria. *Xanthomonas axonopodis* pv. *citri* (*Xac*) causes citrus canker, which affects most commercial citrus cultivars, resulting in significant losses worldwide. Symptoms include canker lesions, leading to abscission of fruit and leaves and general tree decline¹. *Xanthomonas campestris* pv. *campestris*

(*Xcc*) causes black rot, which affects crucifers such as *Brassica* and *Arabidopsis*. Symptoms include marginal leaf chlorosis and darkening of vascular tissue, accompanied by extensive wilting and necrosis². *Xanthomonas campestris* pv. *campestris* is grown commercially to produce the exopolysaccharide xanthan gum, which is used as a viscosifying and stabilizing agent in many industries³. Here we report and compare the complete genome sequences of *Xac* and *Xcc*. Their distinct disease phenotypes and host ranges belie a high degree of similarity at the genomic level. More than 80% of genes are shared, and gene order is conserved along most of their respective chromosomes. We identified several groups of strain-specific genes, and on the basis of these groups we propose mechanisms that may explain the differing host specificities and pathogenic processes.

Xanthomonas axonopodis pv. *citri* (strain 306) has one circular chromosome comprising 5,175,554 base pairs (bp), and two plasmids: pXAC33 (33,699 bp) and pXAC64 (64,920 bp). *Xanthomonas campestris* pv. *campestris* (strain ATCC33913) has a single circular chromosome comprising 5,076,187 bp, and no plasmids (Table 1). The chromosomes of the two organisms have a high degree of collinearity (Fig. 1). The alignment shown in Fig. 1 suggests that just three major rearrangement events have occurred. One is an inversion around the putative terminus of replication (Fig. 1; A); the others (Fig. 1; B1 and B2) are two inversions with translocations symmetrically located with respect to the putative origin of replication. A whole-chromosome alignment at the protein level, excluding transposable elements, paired 2,929 orthologous genes, corresponding to about 70% of the total of each (Fig. 2). There are 512 non-collinear (or translocated) pairs of orthologous genes scattered throughout the two chromosomes. *Xanthomonas campestris* pv. *campestris* contains 646 genes (15.4%) not found in *Xac*, of which 291 have functional assignment. *Xanthomonas axonopodis* pv. *citri* contains 800 genes (18.5%) not found in *Xcc*, of which 316 have functional assignment.

Table 1 General features of *X. campestris* pv. *campestris* and *X. axonopodis* pv. *citri* genomes

(a) General features of the chromosomes		
	<i>Xcc</i>	<i>Xac</i>
Length (bp)	5,076,187	5,175,554
G+C content (%)	65.0	64.7
Protein-coding region (% chromosome size)	84.34	85.59
Protein-coding genes		
With assigned function	2,708	2,710
Conserved hypothetical	1,276	1,272
Hypothetical	198	331
Total	4,182	4,313
Transfer RNA	53	54
Ribosomal RNA operons	2	2
Plasmids	0	2
Insertion sequence elements (IS)	109	87
(b) Sequence similarity between protein-coding genes		
≥80% amino acid identity (%)	82.1	
≥90% amino acid identity (%)	48.9	
≥80% nucleotide identity (%)	78.3	
≥90% nucleotide identity (%)	18.3	
(c) General features of <i>X. axonopodis</i> pv. <i>citri</i> plasmids		
	pXAC33	pXAC64
Length (bp)	33,699	64,920
G+C content (%)	61.9	61.4
Coding region (% plasmid size)	81.59	89.26
Protein-coding genes		
With assigned function	21	39
Conserved hypothetical	7	7
Hypothetical	14	27
Total	42	73
Insertion sequence elements (IS)	11	10

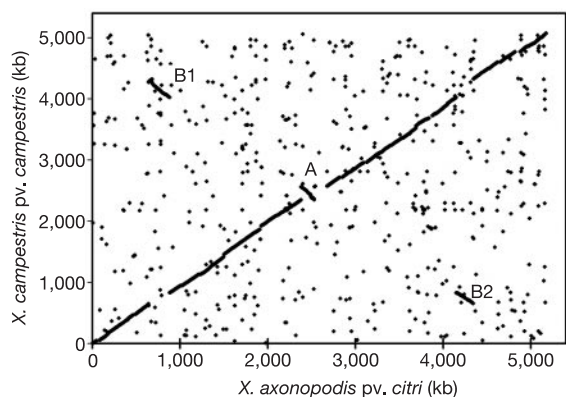


Figure 1 Nucleotide alignment between *Xac* and *Xcc*. Each point is a maximal unique match, that is, a maximal sequence of at least 25 bp that occurs only once in *Xac* and only once in *Xcc*. Each match is exact. Labels A, B1 and B2 are rearrangement events between the genomes.

In both genomes, the respective regions around the putative termini of replication (labelled A in Figs 1 and 2) contain a large number of strain-specific genes, suggesting that these regions are susceptible to gene acquisition and/or gene loss. Notably, some of these genes have functional assignments suggestive of involvement in host colonization and pathogenesis. Genes specific to *Xac* include: (1) *syrE*-related genes, similar to those responsible for the synthesis of the fungicide syringomycin in *Pseudomonas syringae*⁴; (2) a 133,025-bp insertion that harbours many genes not commonly found in γ -proteobacteria and 19 genes also found in the citrus pathogen *Xylella fastidiosa*, including an RTX-like toxin gene with its secretion system (*hlyBD*)⁵; and (3) genes related to the *nodV/nodW* two-component system, which in *Bradyrhizobium japonicum* respond to plant-derived isoflavonone signals and participate in an alternative regulatory pathway for nodulation⁶. Genes specific to *Xcc* include: (1) two copies of the Φ Lf filamentous phage⁷; (2) a gene for cyclic β -1,2-glucan synthetase, associated with virulence/colonization in members of the *Rhizobiaceae* family (*Agrobacterium*, *Rhizobium* and *Brucella*)⁸; (3) the avirulence gene *avrBs1* (ref. 9); (4)

a tannase, which allows microorganisms to overcome growth inhibition induced by plant tannins¹⁰; (5) four genes similar to those involved in antibiotic synthesis in *Streptomyces* (macrolides and streptomycin)¹¹; and (6) a cluster of genes related to nitrate assimilation¹².

The two genomes are rich in transposable elements. We identified 109 genes in *Xcc* and 108 genes in *Xac* that are grouped in 27 different types of transposable element. Only four group types are present in both genomes. In *Xcc*, the IS5 family is highly represented, with 16 copies of IS1478 (ref. 13), whereas in *Xac* the IS3 family is more abundant, with 21 copies of a member not previously described in *Xanthomonas* (ISXac3). Many of these insertion sequences (IS) are located near strain-specific genes, in regions with altered codon usage and G+C content, which suggests that these genes may have been acquired through lateral transfer.

Phylogenetic analyses place *Xac* and *Xcc* at the base of the γ -subdivision of proteobacteria, in the same clade as *X. fastidiosa*. We estimate the divergence time between *Xac* and *Xcc* to be 6.5 to 8.9 million years. Comparison of the predicted *Xac* proteome against other completed genomes (see Supplementary Information) indicates that approximately 40% of all genes found in *Xac* are most similar to genes from non- γ -proteobacteria. Many of these genes have orthologues in the α -proteobacteria *Mesorhizobium loti*, *Sinorhizobium meliloti* and *Agrobacterium tumefaciens*, which have plants as hosts, suggesting that they might be involved in plant-bacterium interactions.

The two *Xanthomonas* pathogens have a large and diversified set of pathways for intermediary, small molecule, and DNA metabolism. The DNA restriction/modification system of *Xac* has one gene cluster for a type I system and another for a type II system, whereas *Xcc* has two clusters for a type I system, but no type II system. This difference may be associated to the noted differences in mobile genetic element content in each genome. *Xanthomonas campestris* pv. *campestris* has a cluster of seven strain-specific genes (*cysG*, *nasTACDEF*) that are required to assimilate and convert nitrate and nitrite into ammonium in *Klebsiella pneumoniae*¹². *Xanthomonas axonopodis* pv. *citri* has strain-specific genes for an ABC-type oligopeptide transport system (*oppA*, *oppB* and *oppC*), which may facilitate the entry of small oligopeptide products for use as a nitrogen source. *Xanthomonas axonopodis* pv. *citri* also contains

Table 2 Proposed *hrpX* regulons

<i>Xanthomonas campestris</i> pv. <i>campestris</i>				<i>Xanthomonas axonopodis</i> pv. <i>citri</i>			
PIP position	Distance (bp)*	Gene ID	Gene product	PIP position	Distance (bp)*	Gene ID	Gene product
<i>hrp</i> gene cluster				<i>hrp</i> gene cluster			
1,434,713	250	XCC1227	HrcQ	475,032	249	XAC0403	HrcQ
1,438,281	70	XCC1230	HrcU	478,627	70	XAC0406	HrcU
1,438,344	80	XCC1231	HrpB1	478,642	81	XAC0407	HrpB1
1,446,618	139	XCC1240	Hpa1	487,556	135	XAC0416	Hpa1
1,446,733	280	XCC1241	Hpa2	488,189	271	XAC0417	Hpa2
Extended Hrp conserved regulon				Extended Hrp conserved regulon			
323,682	65	XCC0258	<i>Xanthomonas</i> conserved hypothetical	332,388	64	XAC0277	<i>Xanthomonas</i> conserved hypothetical
1,962,252	103	XCC1686	Conserved hypothetical	1,966,043	149	XAC1706	Conserved hypothetical
2,172,034	123	XCC1864	β -K-adipate enol-lactone hydrolase	2,191,464	122	XAC1886	β -K-adipate enol-lactone hydrolase
2,666,140	241	XCC2266	Polygalacturonase	2,767,045	179	XAC2374	Polygalacturonase
2,848,531	184	XCC2399	Conserved hypothetical	2,982,678	153	XAC2534	Conserved hypothetical
3,737,359	64	XCC3157	Aminopeptidase	3,886,531	63	XAC3309	Aminopeptidase
4,118,186	118	XCC3459	Endopolygalacturonase	783,305	121	XAC0661	Endopolygalacturonase
Additional Hrp regulon candidate ORFs				Additional Hrp regulon candidate ORFs			
1,010,154	108	XCC0851	Extracellular protease	40,366	78	XAC0337	2-K-3-DdG permease
1,238,085	109	XCC1072	<i>Xanthomonas</i> conserved hypothetical	501,099	107	XAC0424	Conserved hypothetical with GGDEF domain
1,569,371	89	XCC1348	<i>Xanthomonas</i> conserved hypothetical	1,657,406	271	XAC1435	Iron receptor
3,190,239	217	XCC2693	Cysteine protease	2,763,222	227	XAC2370	Conserved hypothetical
				3,805,982	66	XAC3230	Hypothetical protein
				3,428,498	165	XAC2922	HrpW
3,871,764	65	XCC3258	Conserved hypothetical	4,769,953	147	XAC4074	Ribonucleotide-diphosphate reductase
				4,796,299	133	XAC4090	3-oxoacyl-[ACP] reductase

ORF open reading frame.

*Distance (bp) upstream of predicted start codon.

additional strain-specific genes related to proteases, including three genes similar to secreted xanthomonapepsins/pseudomonapepsins. These observations suggest that the two organisms have different nitrate assimilation and oligopeptide absorption capabilities, and this in turn may be linked to their different optimal environments for growth.

Both *Xac* and *Xcc* have an extensive repertoire of genes for cell-wall degradation. Both genomes code for enzymes with cellulolytic, pectinolytic and hemicellulolytic activities (Supplementary Information). *Xanthomonas campestris* pv. *campestris* has more genes involved in pectin and cellulose degradation than *Xac*. In addition, *Xcc* has two 1,4- β -cellobiosidases and two pectin esterases, none of which are found in *Xac*. The lack of these enzymes does not necessarily preclude the degradation of cellulose and pectin by *Xac*. In fact, in citrus canker there is tissue maceration, albeit localized. Black rot, on the other hand, presents massive degeneration of plant tissue. These differences in symptoms may be correlated to the noted differences in genes for cell-wall degradation.

The synthesis of extracellular degrading enzymes and exopolysaccharides is transcriptionally regulated by products of the *rpj* gene cluster. This cluster in *Xcc* has an organization (*rpjABFCHGDI*) similar to that previously described^{14,15}. *Xanthomonas axonopodis* pv. *citri* lacks *rpjH* and *rpjI*, and *X. fastidiosa* also lacks the corresponding genes¹⁶. The expression levels of proteases and endoglucanases are reduced when the *rpjI* gene is inactivated in *Xcc*¹⁵, suggesting that *rpjI* may have a function in the massive tissue degeneration observed in *Xcc* infection. The locus corresponding to *rpjI* in *Xac* is occupied by a truncated copy of an insertion sequence.

Motility in a number of plant pathogenic species is important for infection¹⁷. The genomic sequences of *Xac* and *Xcc* provide a detailed view of the flagellar and chemotactic biosynthetic pathways of *Xanthomonas* species. Chemotaxis receptors and flagella biogenesis genes are organized in four clusters spread over 150 kb in both *Xanthomonas* genomes. One of these clusters has several nearly identical tandemly repeated copies (ten in *Xac* and eight in *Xcc*) of a methyl-accepting chemotaxis protein gene (*mcp*). Such large numbers of tandemly repeated copies are unusual in bacteria, suggesting a prominent role for chemotaxis in the two *Xanthomonas* pathogens. In fact, *Xac* has another ten copies of the *mcp* gene scattered throughout its genome, whereas *Xcc* has eleven additional copies. Both genomes have genes for type IV fimbriae (pili) and glycine-rich outer membrane proteins, which are principal structures for host colonization and adhesion in many pathogenic bacteria¹⁸. However, the fimbrillin genes, whose products may come in direct

contact with host cells or other bacteria, are different in *Xac* and *Xcc*, probably owing to adaptation to specific interactions. Both genomes have copies (two in *Xac*, one in *Xcc*) of a gene that codes for XadA, which is an outer membrane protein found in *Xanthomonas oryzae* pv. *oryzae* (where it is implicated in virulence) and in *Xanthomonas campestris* pv. *vesicatoria* (where it is repressed by *HrpG*¹⁹). The presence of these proteins in two other *Xanthomonas* species may indicate a general importance of these proteins to *Xanthomonas* pathogenicity.

Distinct gene clusters for O-antigen synthesis are present in the genomes of both *Xac* and *Xcc*. They are located in different regions of the respective chromosomes and lack significant sequence similarity. This is consistent with the observation that lipopolysaccharide (LPS) O antigen is pathovar/strain-specific and may be involved in host-range selection and pathogenicity by acting as a barrier against plant toxins²⁰. In *Xcc*, these genes are organized in a single cluster as previously reported²¹, whereas in *Xac*, the genes are divided into two regions. The first region contains genes that code for transferases, epimerases, translocases, and derived sugar transport proteins, and the second contains the *xanAB* and *rmlDABC* genes for nucleotide-sugar and dTDP-L rhamose biosynthesis²¹.

In each genome there are two type II secretion systems that are involved in the secretion of degradative enzymes and toxins. One is the previously described *xps* cluster²², whereas the genes in the second cluster are most similar to genes found in *Caulobacter crescentus*. These clusters are associated with the genomic rearrangement events B1 and B2 (Figs 1 and 2). The type III secretion system (*hrp* pathway) is critical for pathogenicity and initiation of disease. A cluster composed of 26 genes extending from *hpa2* to *hrpF* is found in both genomes. Although these clusters are located in different positions in each genome (Fig. 2), gene order is similar and consistent with the previously characterized *hrp* cluster²³. One exception is *hrpW*, which is located between *hrpF* and *hrpE* in *Xcc*, but outside the *hrp* cluster in *Xac*. The presence of *hrpW* in *Erwinia amylovora*, *Pseudomonas syringae* and now in *Xac* and *Xcc* indicates that *hrpW* is a general core component of plant-pathogen type III systems. The amino acid identity between the orthologous type III genes in *Xac* and *Xcc* is high (greater than 80% on average). The least similar (less than 61% identity) orthologous genes are *HpaP*, *Hpa1*, *HpaA*, *HrpE*, *HrpF* and *HrpW*. It may be significant that some of the proteins coded for by these latter genes are predicted to be exposed components or secreted by the type III secretion system, and differences in these components may reflect different selective pressures due to strain-specific interactions with host plants. Type

Table 3 Putative effector/avirulence genes of *Xcc* and *Xac*

Gene ID	Name	Family	<i>Xcc/Xac</i> *	PIP box†	Location‡	AME§	Motifs
XCC0052/XAC0076	<i>avrBs2</i>	<i>avrBs2</i>	Y/Y	Y/Y	C	N	n
XCC1629/XAC0286	<i>avrXccE1/avrXacE1</i>	<i>avrPphE</i>	Y/Y	Y/Y	C	Y/N	m
XAC3224	<i>avrXacE2</i>	<i>avrPphE</i>	N/Y	N	C	Y	m,n
XACb0011	<i>avrXacE3</i>	<i>avrPphE</i>	N/Y	Y	P	Y	m,n
XACa0022	<i>pthA1</i>	<i>avrBs3</i>	N/Y	N	P	Y	n
XACa0039	<i>pthA2</i>	<i>avrBs3</i>	N/Y	N	P	Y	n
XACb0015	<i>pthA3</i>	<i>avrBs3</i>	N/Y	N	P	Y	n
XACb0065	<i>pthA4</i>	<i>avrBs3</i>	N/Y	N	P	Y	n
XCC2100	<i>avrBs1</i>	<i>avrBs1</i>	Y/N	N	C	Y	n
XCC2099	<i>avrBs1.1</i>	<i>avrBs1</i>	Y/N	N	C	Y	
XCC2109	<i>avrXccC</i>	<i>avrC</i>	Y/N	Y	C	Y	
XCC3731	<i>avrXccB</i>	<i>yopJ</i>	Y/N	Y	C	Y	
XCC4229	<i>avrXccA1</i>	<i>avrXca</i>	Y/N	N	C	N	sp
XCC2396	<i>avrXccA2</i>	<i>avrXca</i>	Y/N	N	C	N	n
XCC4186/XAC3090	Leucine-rich protein	<i>popC</i>	Y/Y	Y/N	C	N/N	lrr
XAC0393	<i>hpaF</i>	<i>popC</i>	Y/N	Y	C	N	lrr
XCC2565	Leucine-rich protein	<i>popC</i>	N/Y	Y	C	N	lrr

* Presence in *Xcc* or *Xac* genome: Y, yes; N, No.

† Presence of PIP motif near promoter: Y, yes; N, No.

‡ Location in genome: C, chromosome; P, plasmid.

§ AME (associated mobile genetic elements): Y, yes; N, No.

|| Amino acid sequence motifs as determined by PSORT or BLAST. m, myristylation; n, nuclear localization; lrr, leucine-rich receptor; sp, bacterial signal peptide cleavage site.

IV secretion systems have been well characterized in organisms such as *Agrobacterium tumefaciens*, where the *virB* operon is required for the transfer of the T-DNA to the plant cell²⁴. Other bacteria use this system for conjugal transfer and secretion of toxins or other proteins²⁴. *Xanthomonas axonopodis* pv. *citri* has two type IV systems, one in the chromosome and one in the pXAC64 plasmid. Gene content in these clusters is different from that of the *Agrobacterium virB* operon: *virB5* and *virB7* are missing from both clusters, and *virD4* is missing in the plasmid. *Xanthomonas campestris* pv. *campestris* has only one *virB* cluster, which is almost identical to that of the *Xac* chromosome, except that *virB6* is outside the cluster. This is the first report to our knowledge of a type IV system in any *Xanthomonas* species, and its precise function remains to be elucidated.

In the genus *Xanthomonas*, expression of structural genes of the type III secretion system and some effector genes are mediated, in part, by the *hrpX* and *hrpG* gene products²⁵, both of which are present in *Xac* and *Xcc*. Several genes that are regulated in a HrpX-dependent manner possess the consensus nucleotide sequence TTCGC...N₁₅...TTCGC, which has been termed the plant-inducible-promoter box, or PIP box²⁶. Detection of a PIP box or a sequence similar to it provides a tool for identification of candidate genes in the *hrpX* regulon as well as effector protein genes of the type III pathway. We found 17 (*Xcc*) and 20 (*Xac*) copies of the sequence TTCG...N₁₆...TTCG (a putative PIP box; separated by 16 nucleotide sequences) in promoter regions (Table 2). Among these, 12 are associated with the same genes in both genomes. Five of these genes are in the *hrp* gene cluster and seven are scattered on the chromosomes, including three with amino-terminal type II signal peptide sequences and similarity to cell-wall-degrading enzymes. Of the five other putative PIP boxes in *Xcc*, two are associated with protease genes. Of the eight other putative PIP boxes in *Xac*, one is associated with an iron receptor gene.

Both genomes have a variety of putative avirulence/effector protein-coding genes, which are fundamental to the development or restriction of the disease (Table 3). *Xanthomonas campestris* pv. *campestris* contains a more diverse array of known types of effectors. However, it lacks *avrBs3/pth*, which is found in *Xac* as well as in many *Xanthomonas* species, and in *Ralstonia solanacearum*²⁷. *Xanthomonas axonopodis* pv. *citri* contains four copies of *avrBs3/pth*, located in the plasmids and without recognizable PIP boxes. These four copies are distinct from each other and from the previously described *pthA*²⁸. *Xanthomonas axonopodis* pv. *citri* also contains three genes with similarity to *avrPphE*, originally identified in *Pseudomonas syringae* pv. *phaseolicola*—*Xcc* has only one *avrPphE* gene. Notably, *Xac* lacks a recognizable YopJ homologue, a protease-coding gene that has been found in many other pathogenic bacteria²⁷ (including animal pathogens). It would be interesting to verify whether the differences noted above contribute significantly to the different infection modes of the two *Xanthomonas* pathogens. Both *Xac* and *Xcc* contain genes coding for leucine-rich-repeat (LRR) proteins. LRR motifs are commonly involved in protein–protein interactions and are found in the three major classes of plant-resistance genes²⁹ and in the PopC protein of *Ralstonia solanacearum*³⁰. The presence of distinct genes for LRR proteins in these genomes, under control of putative PIP boxes, suggests that they may have an important function at some stage of plant colonization.

Comparative analysis between the two *Xanthomonas* strains allowed us to identify a set of strain-specific genes, some of which are probably responsible for the distinct pathogenicity and host specificity profiles of these organisms. *Xanthomonas campestris* pv. *campestris* is uniquely suited to invade and colonize host tissue, a fact that may partially explain the systemic nature of its infection. In contrast, *Xac* induces a strong local response with cell proliferation and necrosis, but shows little spontaneous dissemination, probably due to a smaller number of genes capable of causing a massive

Figure 2 Comparative genome map. The organizations of *Xcc* (upper) and the *Xac* (lower) genomes are represented side by side. Predicted proteins are indicated by vertical bars coloured according to orthologue pairing and collinearity. The genomes were aligned by determining a longest common subsequence (LCS) of orthologues, with manual adjustments. Light blue, collinear orthologous proteins; dark blue, translocated (that is, not belonging to the LCS) orthologous proteins; yellow, *Xcc*-specific proteins; pink, *Xac*-specific proteins. Green flags indicate transposases. Rearrangement events are labelled A, B1 and B2, and are indicated by red lines. Numbers in black represent genome coordinates in base pairs. Features in each genome are labelled as follows: O antigen, O-antigen biosynthesis cluster; Xcs and Xps, secretory pathway II; Hrp, secretory pathway III; SyrE, syringomicin-related proteins; GS, cyclic β -1,2-glucan synthetase; Tan, tannase protein; RTX, haemolysin calcium-binding protein; RM.I or .II, restriction and modification system type I or II; XccP1 or XacP1, Φ CTX-type prophage; Vir, VirB proteins; Lf, Φ Lf prophage; RpfI, RpfI protein; RpfH, RpfH protein; Nas, Nitrate assimilation proteins; Opp, oligopeptide transport system; M/S, macrolide/streptomycin-related proteins; VW, NodV and NodW proteins.

degeneration of host tissue. Besides the already known *pthA* genes, other candidates involved in eliciting this strong host response should emerge experimentally by testing the roles of a selected list of *Xac*-specific genes currently without functional assignment. Host specificity may result from a combination of differential subsets of genes found in each genome, such as *avr*, type III secretion system, *rpf*, type IV fimbriae, and LPS O-antigen operons. Future experimental tests of the several hypotheses put forward here will be facilitated now that the genomes of a plant pathogen (*Xcc*) and one of its hosts (*Arabidopsis thaliana*) are publicly available. □

Methods

The sequencing and analysis in this work were carried out by the *Xanthomonas* consortium of the Organization for Nucleotide Sequencing and Analysis (ONSA), a group of 13 biology laboratories and one bioinformatics centre, located in the state of São Paulo, Brazil.

Sequencing and assembly

We sequenced both genomes using shotgun methodologies. Several pUC18 libraries were generated with inserts ranging from 1 to 4 kb in size, and sequenced from both ends. From these libraries 205,408 (*Xac*) and 174,051 (*Xcc*) reads were sequenced. Several cosmid libraries (using lawlrit vector) were generated, with inserts ranging from 30 to 45 kb in size. Both ends of these inserts were sequenced, resulting in 7,380 (*Xac*) and 5,476 (*Xcc*) ends. Assembly was carried out using the phred/phrap/consed package and a scaffolding program developed by our bioinformatics team. Assembly was confirmed by comparing *CeuI*, *SwaI* and *PmeI* restriction maps to computational predictions. Genome assembly yielded 103 virtual gaps in *Xac* and 135 in *Xcc*. Many of these gaps were a result of the repeat screening process required for scaffolding. Gaps were closed by whole-plasmid and cosmid sequencing, plus refill of screened sequences. The total number of whole plasmids sequenced was 20 (*Xac*) and 67 (*Xcc*). The total number of whole cosmids sequenced was 49 (*Xac*) and 13 (*Xcc*). The final *Xac* sequence has an estimated overall error rate of less than 1 in 10,000 bp, and no base has phred error quality below 20. The *Xcc* sequence has 15 uncertain bases due to resolution errors within regions of secondary structure.

Putative origins of replication

On the basis of similarity with *X. fastidiosa*, we hypothesize that the origins of replication for both chromosomes are located in a region between the 50S ribosomal protein L34 gene and the *gyrB* gene. This is supported by predictions based on GC-skew inversion. The origins of the *Xac* plasmids were chosen according to GC-skew inversion and within a region that did not contain predicted genes.

Genome annotation

Annotation was carried out using a system developed by our bioinformatics team. Putative protein-coding genes were identified using GeneMark and Glimmer. Curators assigned functions to these genes based on comparisons to sequences available in public databases. RNA species were identified using BLASTN and tRNAscan-SE. Metabolic pathways were analysed using the KEGG website (<http://www.jaist.genome.ad.jp/kegg>). Transporter proteins were annotated based on BLAST comparisons of predicted proteins in each proteome against a custom database of transporters, which was created based on data available from <http://www-biology.ucsd.edu/~msaier/transport/toc.html#1A>.

Phylogeny

Phylogenetic analyses were performed based on 16S ribosomal RNA sequences using the program DNADIST from PHYLIP, with a distance matrix based on the two-parameter Kimura model. Divergence time was estimated based on the 16S rRNA sequences. A constant rate of 1–2% per 50 million years was assumed.





Whole-genome comparisons between *Xac* and *Xcc*

We compared the genomes at the nucleotide level using the program MUMmer with default values. At the amino acid level the genomes were compared using programs developed by our bioinformatics team. Genes *g* and *h* were considered orthologues if *h* is the best BLASTP hit for *g* and vice versa, with *e*-values less than or equal to 10^{-20} . A gene was considered strain-specific if it had no hits with an *e*-value 10^{-5} or less.

Received 11 December 2001; accepted 26 March 2002.

1. Gottwald, T. R. & Graham, J. H. in *Compendium of Citrus Diseases* (eds Timmer, L. W., Garnsey, S. M. & Graham, J. H.) 5–7 (Am. Phytopathol. Soc., St. Paul, 2000).
2. Hayward, A. C. in *Xanthomonas* (eds Swings, J. G. & Civerolo, E. L.) 51–54 (Chapman & Hall, London, 1993).
3. Becker, A., Katzen, F., Puhler, A. & Ielpi, L. Xanthan gum biosynthesis and application: a biochemical/genetic perspective. *Appl. Microbiol. Biotechnol.* **50**, 145–152 (1998).
4. Menestrina, G. & Semjen, B. V. in *The Comprehensive Sourcebook of Bacterial Proteins Toxins* (ed. Alouf, J. E.) 287–309 (Academic, London, 1999).
5. Simpson, A. J. *et al.* The genome sequence of the plant pathogen *Xylella fastidiosa*. *Nature* **406**, 151–157 (2000).
6. Göttfert, M., Grob, P. & Henneke, H. Proposed regulatory pathway encoded by the *nodV* and *nodW* genes, determinants of host specificity in *Bradyrhizobium japonicum*. *Proc. Natl Acad. Sci. USA* **87**, 2680–2684 (1990).
7. Chang, K. H. *et al.* Sequence analysis and expression of the filamentous phage phi Lf gene I encoding a 48-kDa protein associated with host cell membrane. *Biochem. Biophys. Res. Commun.* **245**, 313–318 (1998).
8. Breedveld, M. W., Hadley, J. A. & Miller, K. J. A novel cyclic β -1,2-glucan mutant of *Rhizobium meliloti*. *J. Bacteriol.* **177**, 6346–6351 (1995).
9. Ronald, P. C. & Staskawicz, B. J. The avirulence gene *avrBs1* from *Xanthomonas campestris* pv. *vesicatoria* encodes a 50-kD protein. *Mol. Plant-Microbe Interact.* **1**, 191–198 (1988).
10. Bhat, T. K., Singh, B. & Sharma, O. P. Microbial degradation of tannins—a current perspective. *Biodegradation* **9**, 343–357 (1998).
11. Beyer, S., Distler, J. & Piepersberg, W. The *str* gene cluster for the biosynthesis of 5'-hydroxystreptomycin in *Streptomyces glaucescens* GLA.0 (ETH 22794): new operons and evidence for pathway-specific regulation by *StrR*. *Mol. Gen. Genet.* **250**, 775–784 (1996).
12. Lin, J. T., Goldman, B. S. & Stewart, V. The *nasFEDCBA* operon for nitrate and nitrite assimilation in *Klebsiella pneumoniae* M5a1. *J. Bacteriol.* **176**, 2551–2559 (1994).
13. Chen, J. H., Hsieh, Y. Y., Hsiao, S. L., Lo, T. C. & Shau, C. C. Characterization of insertions of IS476 and two newly identified insertion sequences, IS1478 and IS1479, in *Xanthomonas campestris* pv. *campestris*. *J. Bacteriol.* **181**, 1220–1228 (1999).
14. Tang, J. L. *et al.* Genetic and molecular analysis of a cluster of *rpf* genes involved in positive regulation of synthesis of extracellular enzymes and polysaccharide in *Xanthomonas campestris* pathovar *campestris*. *Mol. Gen. Genet.* **226**, 409–417 (1991).
15. Dow, J. M., Feng, J. X., Barber, C. E., Tang, J. L. & Daniels, M. J. Novel genes involved in the regulation of pathogenicity factor production within the *rpf* gene cluster of *Xanthomonas campestris*. *Microbiology* **146**, 885–891 (2000).
16. Dow, J. M. & Daniels, M. J. *Xylella* genomics and bacterial pathogenicity to plants. *Yeast* **17**, 263–271 (2000).
17. Rudolph, K. in *Xanthomonas* (eds Swings, J. G. & Civerolo, E. L.) 193–264 (Chapman & Hall, London, 1993).
18. Cao, H., Baldini, R. L. & Rahme, L. G. Common mechanisms for pathogens of plants and animals. *Annu. Rev. Phytopathol.* **39**, 259–284 (2001).
19. Noël, L., Thieme, E., Nennstiel, D. & Bonas, U. cDNA-AFLP analysis unravels a genome-wide *hrpG*-regulon in the plant pathogen *Xanthomonas campestris* pv. *vesicatoria*. *Mol. Microbiol.* **41**, 1271–1281 (2001).
20. Dow, J. M., Osbourn, A. E., Wilson, T. J. & Daniels, M. J. A locus determining pathogenicity of *Xanthomonas campestris* is involved in lipopolysaccharide biosynthesis. *Mol. Plant-Microbe Interact.* **8**, 768–777 (1995).
21. Vorhölter, F. J., Niehaus, K. & Puhler, A. Lipopolysaccharide biosynthesis in *Xanthomonas campestris* pv. *campestris*: a cluster of 15 genes is involved in the biosynthesis of the LPS O-antigen and the LPS core. *Mol. Genet. Genomics* **266**, 79–95 (2001).
22. Dums, F., Dow, J. M. & Daniels, M. J. Structural characterization of protein secretion genes of the bacterial phytopathogen *Xanthomonas campestris* pathovar *campestris*: relatedness to secretion systems of other gram-negative bacteria. *Mol. Gen. Genet.* **229**, 357–364 (1991).
23. Bonas, U. *et al.* Isolation of a gene cluster from *Xanthomonas campestris* pv. *vesicatoria* that determines pathogenicity and the hypersensitive response on pepper and tomato. *Mol. Plant-Microbe Interact.* **4**, 81–88 (1991).
24. Burns, D. L. Biochemistry of type IV secretion. *Curr. Opin. Microbiol.* **2**, 25–29 (1999).
25. Wengelnik, K. & Bonas, U. HrpXv, an AraC-type regulator, activates expression of five of the six loci in the *hrp* cluster of *Xanthomonas campestris* pv. *vesicatoria*. *J. Bacteriol.* **178**, 3462–3469 (1996).
26. Fenselau, S. & Bonas, U. Sequence and expression analysis of the *hrpB* pathogenicity operon of *Xanthomonas campestris* pv. *vesicatoria* which encodes eight proteins with similarity to components of the Hrp, Ysc, Spa, and Fli secretion systems. *Mol. Plant-Microbe Interact.* **8**, 845–854 (1995).
27. Lahaye, T. & Bonas, U. Molecular secrets of bacterial type III effector proteins. *Trends Plant Sci.* **6**, 479–485 (2001).
28. Yang, Y. & Gabriel, D. W. Intragenic recombination of a single plant pathogen gene provides a mechanism for the evolution of new host specificities. *J. Bacteriol.* **177**, 4963–4968 (1995).
29. Young, N. D. The genetic architecture of resistance. *Curr. Opin. Plant Biol.* **3**, 285–290 (2000).
30. Gueneron, M., Timmers, A. C., Boucher, C. & Arlat, M. Two novel proteins, PopB, which has functional nuclear localization signals, and PopC, which has a large leucine-rich repeat domain, are secreted through the hrp-secretion apparatus of *Ralstonia solanacearum*. *Mol. Microbiol.* **36**, 261–277 (2000).

Supplementary Information accompanies the paper on Nature's website (<http://www.nature.com>).

Acknowledgements

We thank A. M. da Silva, S. Verjovski-Almeida and D. W. Wood for reading of the manuscript; our steering committee (S. Oliver, A. Goffeau, J. Sgouros, A. C. M. Paiva and J. L. Azevedo) for their critical accompaniment; and all the technicians in the sequencing laboratories of ONSA involved in this project. Project funding was from FAPESP, Fundecitrus, Fundect-MS and CNPq.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to A.C.R.d.S. (e-mail: acrasera@iq.usp.br). The sequences have been deposited in GenBank under accession numbers AE008922 (*Xcc* chromosome), AE008923 (*Xac* chromosome), AE008924 (pXAC33) and AE008925 (pXAC64).

The catalytic pathway of horseradish peroxidase at high resolution

Gunnar I. Berglund^{*†}, Gunilla H. Carlsson^{*†}, Andrew T. Smith[‡], Hanna Szöke^{*§}, Anette Henriksen^{§||} & Janos Hajdu^{*}

^{*} Department of Biochemistry, Uppsala University, Biomedical Center, Box 576, S-751 23 Uppsala, Sweden

[‡] School of Biological Sciences, University of Sussex, Brighton BN1 9QG, UK
^{||} Protein Structure Group, Department of Chemistry, University of Copenhagen, Universitetsparken 5, DK-2100 København, Denmark

[†] These authors contributed equally to this work

A molecular description of oxygen and peroxide activation in biological systems is difficult, because electrons liberated during X-ray data collection reduce the active centres of redox enzymes catalysing these reactions^{1–5}. Here we describe an effective strategy to obtain crystal structures for high-valency redox intermediates and present a three-dimensional movie of the X-ray-driven catalytic reduction of a bound dioxygen species in horseradish peroxidase (HRP). We also describe separate experiments in which high-resolution structures could be obtained for all five oxidation states of HRP, showing such structures with preserved redox states for the first time.

In 1810, Planche⁶ reported that a tincture of guaiacum developed a stronger colour when a piece of fresh horseradish root was soaked in it. During the next two centuries, horseradish peroxidase (HRP), the haem-containing enzyme responsible for the colour change in Planche's experiment, was used extensively to develop ideas on redox catalysis. Haem-containing redox enzymes participate in a strikingly diverse range of chemistry, yet all biological oxidation reactions catalysed by these enzymes involve very similar high-oxidation-state intermediates (Fig. 1) whose reactivity is modulated by the protein environment^{4,5,7,8}. An understanding of these enzymes in structural terms requires detailed structures for the redox intermediates. This is not a trivial task because electrons liberated in the sample by X-rays during crystallographic data collection alter the redox state of the active site^{1–5}. Redox enzymes have evolved to channel electrons efficiently into an oxidized active site. Attempts to minimize radiation-induced structural changes in crystals^{1,2,4,9–11} include soaking them in excess electron scavengers¹² and using short-wavelength X-rays⁴. Neither of these measures proved to be sufficient to preserve the redox state of oxidized intermediates in HRP during conventional X-ray data collection. Figure 2a and b shows that the ferric and the compound III forms of HRP became reduced by X-rays at doses smaller than those required

§ Present addresses: Lawrence Livermore National Laboratory, PO Box 808, L-41, Livermore, California 94551, USA (H.S.); Carlsberg Laboratory, Gamle Carlsberg Vej 10, DK-2100 Valby, Denmark (A.H.).

for the collection of a complete diffraction data set (wavelength range tested, 0.78–0.98 Å). These findings have implications beyond the immediate scope of this paper and raise questions about the actual oxidation state of many redox proteins in the Protein Data Bank. There is an intrinsic limit to the amount of data which can be extracted from a sample of a given size^{13–15} in a given time^{16–18}. Figure 2c illustrates the difficulty in obtaining a structure for the bound dioxygen species in compound III. At the start of data collection, the population of compound III was nearly 100% in the crystal (Fig. 2b). However, the electron density map in Fig. 2c gives no sign of the bound dioxygen species; instead, densities for separated water molecules appear in the active site. Electrons liberated in the sample and funnelled to the active site seem to have been used by the enzyme to cleave the bound dioxygen species. We interpret the map of Fig. 2c as a mixture of structures dominated by the products.

In order to minimize such side reactions, we employed a multi-crystal data collection strategy based on a systematic spread of the X-ray dose over many crystals (Fig. 3a). Single-crystal microspectrophotometry was then used to correlate measured structural transitions with measured electronic transitions in experiments similar to stoichiometric redox titrations. This approach permits one to drive (and control) redox reactions in crystals, using a small number of electrons redistributed in the sample crystals (for further details see Supplementary Information). Using this technique, we captured compound III (Fig. 3b top) and obtained a three-dimensional movie on the X-ray-driven catalytic conversion of the bound dioxygen species to two molecules of water in the active site (Fig. 3b). The electronic structure of compound III of HRP is similar to that of oxyhaemoglobin and oxyhaemoglobin and can be described as $\text{Fe}^{2+}\text{-O}_2$ or the isoelectronic $\text{Fe}^{3+}\text{-O}_2^{\cdot-}$ form¹⁹. The dioxygen bound to HRP has a significant superoxide character²⁰ and is highly reactive. It offers an excellent model on which the four-electron reduction of dioxygen to water can be studied in structural terms. The atomic details of this reaction have been elusive, owing to difficulties in obtaining reliable X-ray structures for oxidized redox intermediates. Figure 3b shows frames assembled from small chunks of data as shown in Fig. 3a. The results show the breakage of the oxygen–oxygen bond, the release of the first water molecule, formation of a ‘ferryl-like’ intermediate, and the formation of the second water molecule in agreement with the expected stoichiometry of the reaction. Electron transfer reactions and proton-coupled electron transfer processes readily occur at cryogenic temperatures in protein crystals along with small-scale movements of ligands and side chains^{4,5,21–23}. The water molecule derived from the distal oxygen atom (and formed first) is within hydrogen-bonding distance of the oxygen atom remaining on the iron. This water also forms hydrogen bonds with the Nε2 atom of His 42, the side-chain of Arg 38 and another water peak (not shown here for the

sake of clarity, but shown in the high-resolution structures of compounds I and II captured in separate experiments; Fig. 4d, e). At higher X-ray doses, the iron–oxygen bond breaks, leaving behind a penta-coordinated iron, which is now 0.2 Å below the plane of the pyrrole nitrogens with the iron in the ferrous state. Structural changes were limited to the active site with no measurable changes in the structures of disulphide bridges, suggesting that the active site is significantly more sensitive to reduction than the protein itself. The sequence represents changes in the average structure at the active site during exposure to X-rays. Figure 3c shows a mechanism

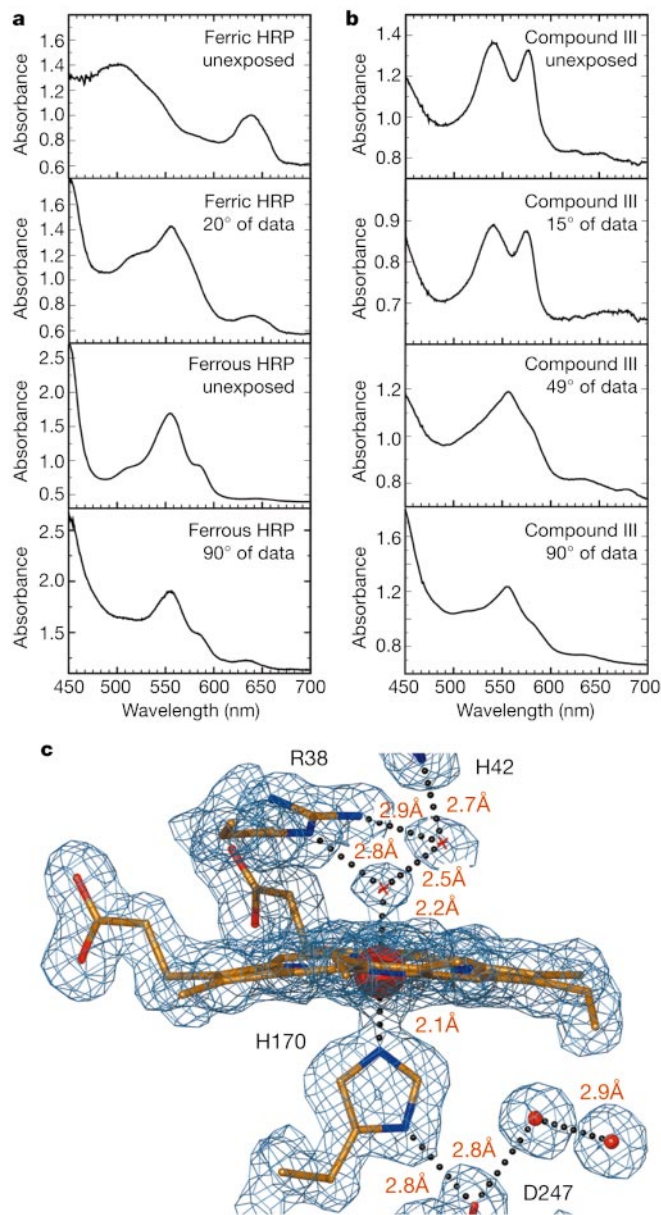


Figure 2 X-ray-induced reduction of horseradish peroxidase (wavelength range, 0.93–0.98 Å). **a**, Changes in the spectra of the ferric and ferrous forms following X-ray exposure. **b**, Spectral changes in compound III. **c**, The 1.60-Å electron-density map calculated from 90 degrees of data from a single compound III crystal. Accession code, PDB 1h5m. The density shows two solvent molecules above the iron instead of the expected dioxygen species. SigmaA-weighted³⁰ $2mF_{\text{obs}} - DF_{\text{calc}}$ maps contoured at 1σ (σ represents the root mean square electron density for the unit cell). Carbons are coloured gold, nitrogens blue and oxygens red. This colouring scheme is used in all other figures.

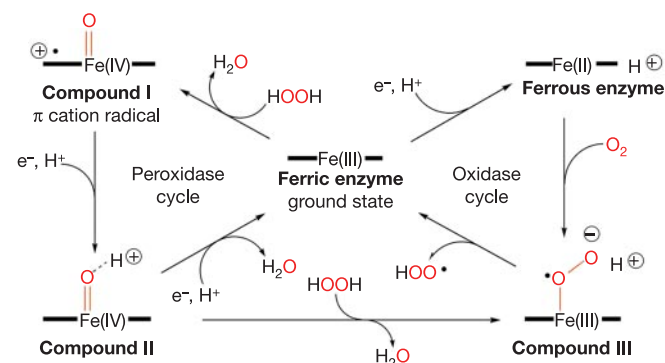


Figure 1 The five oxidation states of horseradish peroxidase.

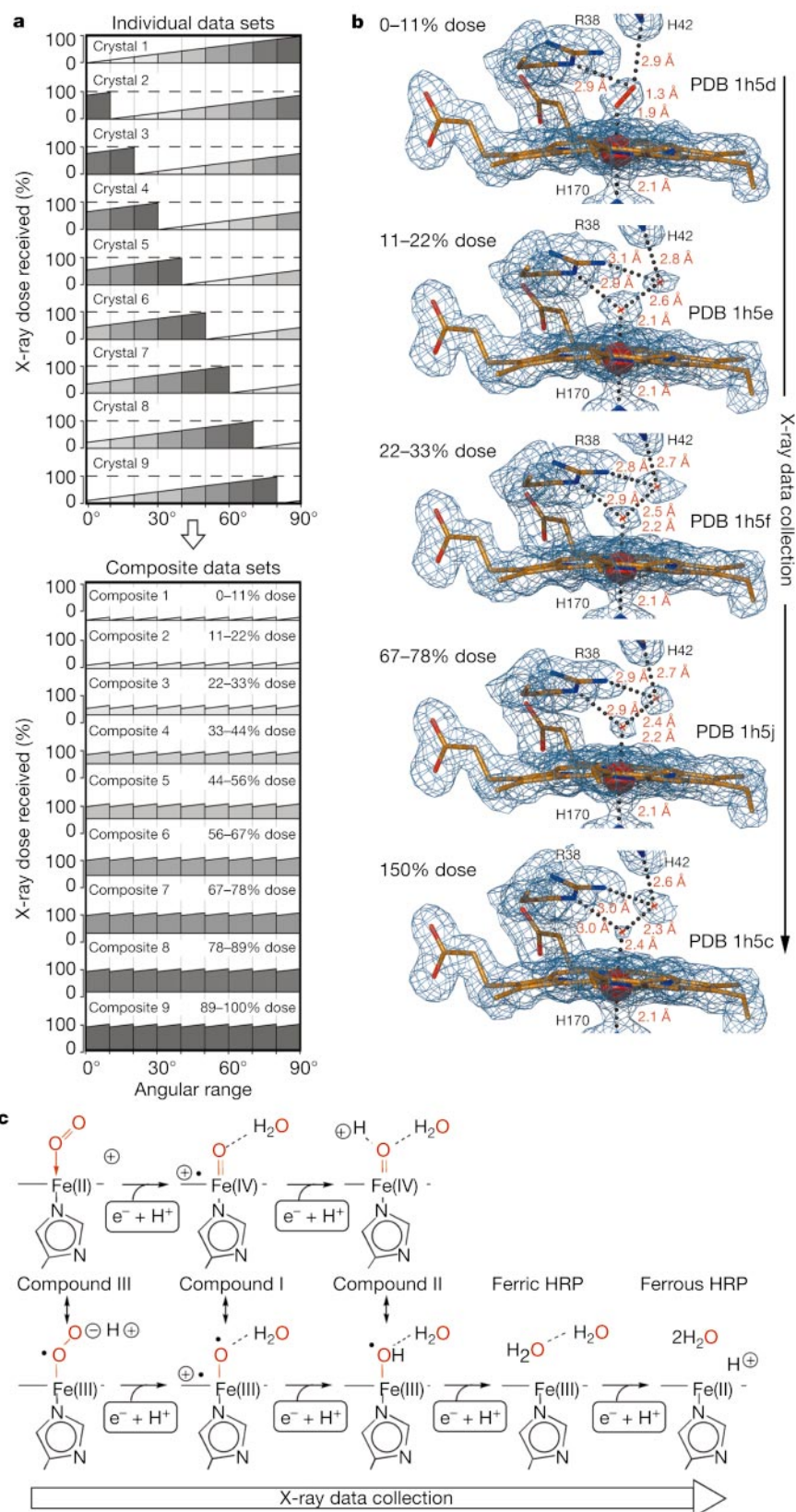


Figure 3 X-ray-driven catalytic conversion of a dioxygen species in horseradish peroxidase. **a**, The multocrystal data collection strategy, showing the distribution of the X-ray dose as a function of the rotation angle on individual (and spectrally uniform) crystals of HRP. The construction of composite data sets from small chunks of the individual data sets is shown at the bottom. Composite data sets represent structures that received different X-ray doses. This method permits experiments similar to redox titrations.

b, SigmaA-weighted³⁰ $2mF_{\text{obs}} - DF_{\text{calc}}$ maps contoured at 1σ showing X-ray-induced reduction of compound III. For the last structure, the crystal was pre-exposed to X-rays for 90° before another full X-ray data set was collected on it. Accession codes are shown. **c**, A possible mechanism for the reduction of the bound dioxygen species to two molecules of water. Structures linked by double arrows are isoelectronic with each other.

which can account for the X-ray-induced transformation of the bound dioxygen species of compound III to two molecules of water in HRP.

We investigated whether this is all that happens in terms of movements during catalysis in the crystal. We performed a second set of experiments where each of the five expected intermediates was built up to high concentrations in crystals of unfrozen HRP. Each of these intermediates was then preserved for subsequent structure determination by quenching the crystals in liquid nitrogen (see Methods). Structures obtained from the first few degrees of data for these intermediates are shown in Fig. 4 and the crystallographic data are summarized in the Supplementary Information.

Figure 4a shows the structure of the pure ferric form. The ferric iron is penta-coordinated and is predominantly high-spin^{7,8}. Density for an acetate ion (from the crystallization solution, see Methods) is visible above the haem plane. In the structure of the acetate-free form of the enzyme²⁴, two solvent molecules take up the oxygen positions of the acetate. Acetate can be washed away from the crystal and is not present in the structures of intermediates (Fig. 4b–e).

Figure 4b shows the structure of the pure ferrous form. This structure was obtained by soaking ferric crystals in dithionite before freezing the crystal in liquid nitrogen. The enzyme remained in the reduced, ferrous state during data collection, and a full data set

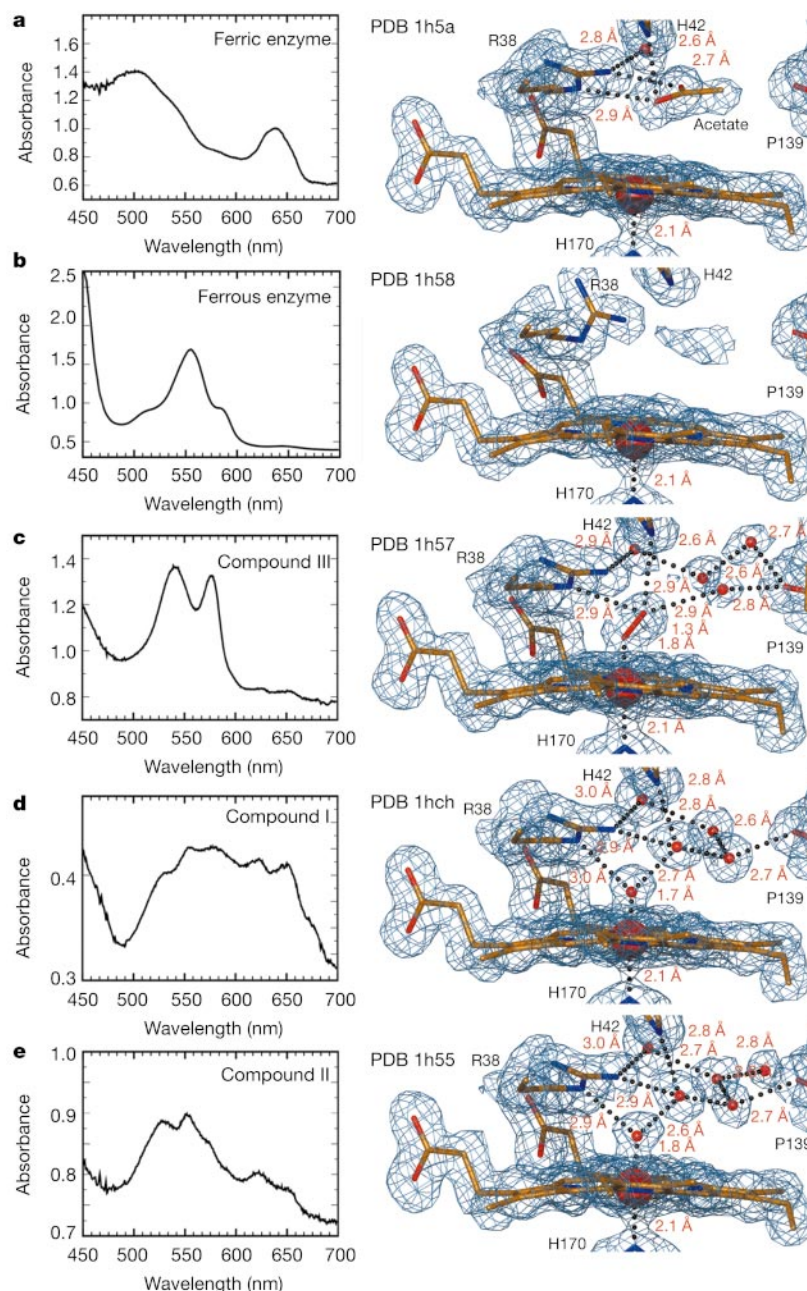


Figure 4 Spectra and refined high-resolution structures for the five oxidation states of horseradish peroxidase. Accession codes are shown. The structures were captured at high concentrations in individual experiments (Methods) and were obtained from the first few degrees of data (Table 2 of the Supplementary Information). **a**, Ferric enzyme.

b, Ferrous enzyme. **c**, Compound III (Fe–O bond distance, 1.8 Å). **d**, Compound I (Fe–O bond distance, 1.7 Å). **e**, Compound II (Fe–O bond distance, 1.8 Å). SigmaA-weighted³⁰ $2mF_{\text{obs}} - DF_{\text{calc}}$ maps contoured at 1σ .

could be collected on one crystal (Fig. 2a, bottom).

Figure 4d shows the structure of an 80% pure compound I in which one electron is withdrawn from the ferric iron and a second from the porphyrin ring to give an oxoferryl species and a porphyrin π -cationic radical. Compound I was produced by reacting ferric HRP crystals with peracetic acid. Peracetic acid can oxidize HRP to compound I, but cannot react with it further; thus a practically complete conversion of the crystal to compound I can be achieved. In order to suppress unwanted side reactions, we used crystals grown from trans-3-methoxycinnamic acid, a compound which is not susceptible to oxidation by the enzyme. At the beginning of data collection, the visible absorption spectrum of the crystals suggested nearly pure compound I. Compound I is, however, extremely labile and about a third of it was converted to compound II already after the collection of 8° of data at 0.78 Å wavelength. The structure calculated from composite data obtained from 11 crystals (Fig. 4d) is predominantly compound I (80%), but it also contains about 20% compound II on average, as judged by the spectral data. The iron–oxygen bond distance in this structure is 1.7 Å, which agrees well with extended X-ray absorption fine structure (EXAFS) measurements (1.64 Å)²⁵. Electron-density maps calculated with the program EDEN²⁶ (Methods) confirm our interpretation of the data. The refined structure indicates an occupancy of at least 85% for the bound oxygen atom on the iron. This value is based on the assumption that the *B*-factor of the covalently attached oxygen atom was not significantly different from *B*-factors in its immediate vicinity. The ferryl oxygen forms hydrogen bonds with the Nε atom of Arg 38 (3.0 Å) and with a water molecule (2.7 Å). This water is hydrogen-bonded to His 42 (2.8 Å) and Arg 38 (2.9 Å, Fig. 4d) and occupies the position of the first-formed water in Fig. 3.

Figure 4e shows the structure of compound II. Spectral data show that this species was at least 80% pure. The ferryl oxygen of compound II has an occupancy close to 100%. There is no direct hydrogen bond between His 42 and the ferryl oxygen in compounds I or II. The ferryl oxygen of compound II makes hydrogen bonds to the Nε atom of Arg 38 (2.9 Å) and to a water molecule (2.6 Å) hydrogen-bonded to His 42 (2.8 Å) and Arg 38 (2.9 Å, Fig. 4e). The active site of HRP stabilizes a net positive charge on the haem in compound I, but in compound II this π -cationic radical is no longer present. The electron densities in Fig. 4d and e show a hydrogen bond between the ferryl oxygen and the density at the position of the first-formed water molecule in Fig. 3. This interaction is very similar in compounds I and II. In compound II, however, a proton is believed to have moved to the ferryl oxygen, probably from His 42 via the intervening water-molecule, priming the formation of the second product water in the reaction. The iron–oxygen bond is 0.14 Å longer in compound II than in compound I, giving an iron–oxygen distance of 1.8 Å (1.93 Å by EXAFS²⁵). The longer Fe–O bond supports the argument for the migration of a proton to the ferryl oxygen, following the removal of the π -cationic radical in compound II.

Figure 4c shows the structure of a 95% pure compound III. The refined occupancy of the dioxygen species was about 95%. The dioxygen species ligates the haem iron in a bent conformation with the OA atom 1.8 Å away from the iron (1.72 Å by EXAFS²⁵). The Fe–OA–OB angle is 126°. The OB atom forms hydrogen bonds with the Nε2 atom of His 42 (2.9 Å), the Nε atom of Arg 38 (2.9 Å) and with a water molecule (2.9 Å). His 42 moves 0.5 Å towards the distal oxygen atom of the bound dioxygen species. Arg 38 moves 0.2 Å towards this oxygen and forms a hydrogen bond. The haem iron together with His 170 moves 0.2 Å towards the plane of the pyrrole nitrogens as a result of oxygen binding.

All structural results show changes specific to the active site. There are four disulphide bridges in HRP and they remained unperturbed during data collection on compounds I, II, III and on the ferric enzyme. The Cys 97–Cys 301 disulphide bridge is partially broken in the ferrous enzyme, probably as a result of

reduction with dithionite. Cys 11–Cys 91 can be modelled in two conformations in all structures, which suggests genuine disorder in this area in HRP.

Charge-exchange reactions between highly oxidized metal centres and neutral molecules are important in difficult synthetic and degradative processes in chemistry and biology and have recently been identified as the source of cometary X-ray emission²⁷. After dioxygen began to be produced by photosynthetic organisms on Earth, the maximum oxidation/ionization states attainable by metal centres in biocatalysis rose swiftly. The catalytic intermediates of HRP described in this paper show these structures with preserved redox state for what we believe is the first time. We suggest that similar structures exist in cytochromes P450 and the terminal oxidases of aerobic respiratory chains. The results presented here provide direct structural insights into the redox chemistry behind these processes at porphyrin-based iron centres. □

Methods

Crystallization of ferric HRP

Crystals²⁸ of recombinant HRP²⁹ were grown from 20% polyethylene glycol (PEG) 8000, 0.2 M calcium acetate, 0.1 M sodium cacodylate, pH 6.5 by vapour diffusion at 4 °C, using streak seeding. Two crystallizations were set up: (1) For studies on the ferric, ferrous, compound II and compound III forms, the droplets contained 2 µl protein (7.3 mg ml^{−1} in 0.1 M cacodylate, pH 6.5), 2 µl reservoir solution and 1 µl isopropanol saturated with ferulic acid. Before the experiments, crystals were washed twice overnight in 5 µM hydrogen peroxide in reservoir solution to remove bound ferulic acid. (2) For studies on compound I, crystals were grown in droplets containing 3 ml protein (7.3 mg ml^{−1}), 3 µl reservoir solution and 0.5 µl isopropanol saturated with trans-3-methoxycinnamic acid. Trans-3-methoxycinnamic acid was removed by washing these crystals in reservoir solution twice overnight. Absorption spectra showed that both types of washed crystals were in the ferric state. Consequent reactions in non-frozen crystals were all performed at 4 °C.

Preparation of ferrous crystals

The ferrous state was formed by soaking ferric crystals for 30 min in a reservoir solution supplemented with 100 mM sodium dithionite and 10% PEG 400.

Preparation of compound III in crystals

Compound III developed within 2–5 min after placing ferric HRP crystals in 54 mM hydrogen peroxide, 10 µM ferulic acid, and 10% PEG 400.

Preparation of compound I in crystals

Compound I was produced by soaking washed crystals in a droplet of 1–2 mM peracetic acid dissolved in reservoir solution. Compound I has a greenish colour that developed in 2–10 min. Crystals were flash frozen in liquid nitrogen, following exposure to 10% PEG 400 in reservoir solution for cryoprotection.

Preparation of compound II in crystals

Compound II was produced in washed crystals following a soak in a droplet of 2–5 mM peracetic acid dissolved in reservoir solution, which also contained 10% PEG 400. After the green colour of compound I had disappeared (2–20 min), the crystals were flash frozen.

Flash freezing and single crystal microspectrophotometry

All crystals were flash frozen in liquid nitrogen and stored for data collection. Their spectra were measured at 100 K before and after X-ray exposure with a microspectrophotometer (4DXray Systems AB; <http://www.4dx.se>) linked to an Oxford CryoStream cooler (OxfordCryosystems; <http://www.oxfordcryosystems.co.uk>). Wherever possible, care was taken to record spectra in similar crystal orientations. Spectral deconvolution was performed with SigmaPlot 3.0 (Jandel Scientific). Spectra of frozen crystals remained unchanged during storage.

X-ray data collection and processing

Data were collected at 100 K at BM14 and ID14 EH3 of the ESRF (Grenoble, France) and at I-711 at MaxLab (Lund, Sweden) with X-rays of the shortest possible wavelengths. Crystals of HRP were aligned to allow collection of a complete data set with 90 degrees rotation. The strategies outlined in Fig. 3a were applied. Data were processed and scaled using the HKL package (<http://www.hkl-xray.com>) and programs in the CCP4 suite (<http://www.ccp4.ac.uk>).

Crystallographic refinement

Starting coordinates (7atj; ref. 28) included the protein moiety and the haem group. Identical test sets (with reflections of the same indices) were used to calculate and compare *R*_{free} values for the different structures. Models were refined with an amplitude-based maximum-likelihood target in the program CNS (<http://cns.csb.yale.edu>) applying overall *B*-factor corrections and bulk solvent corrections. In the final stages of the refinements, the iron, the porphyrin and any ligand bound to the distal side of the haem iron were refined as

separate non-bonded groups using only X-ray terms in the target function. SigmaA-weighted electron-density maps³⁰ and the program EDEN²⁶ were used to check and to eliminate model bias. In the EDEN runs, the haem and its ligands were removed from the protein model and EDEN was used with a gentle solvent target to recover these missing molecules completely. In all cases shown here, the haem and its bound ligands were recovered and the electron density was at least as clear as in the starting maps, indicating that refinement proceeded without model bias. In order to perform further tests, the EDEN solution was perturbed by independently changing amplitudes and phases in the file of the calculated structure factor amplitudes (F_{calc}), using a 30% random gaussian perturbation; the resulting perturbed F_{calc} file served as the starting model for another EDEN run. The perturbation/recovery steps were repeated 10 times and the resulting 10 maps were then averaged. These maps showed excellent agreements with the starting maps. Occupancy of the ligands were refined in SHELXL (<http://shelx.uni-ac.gwdg.de/SHELXL>).

Models were inspected with O (<http://xray.bmc.uu.se/alwyn>) and figures were rendered by POV-Ray (<http://www.povray.org>), using the Molray interface at <http://xray.bmc.uu.se/markh>.

Coordinates and structure factors have been deposited in the Protein Data Bank (accession codes 1h5m, 1h5d, 1h5e, 1h5f, 1h5g, 1h5h, 1h5i, 1h5j, 1h5k, 1h5l, for the structures in Table 1 of the Supplementary Information and 1h58, 1h5a, 1h57, 1h5c, 1hch and 1h55 for the structures in Table 2 of the Supplementary Information).

Received 11 November 2001; accepted 4 March 2002.

1. Chance, B., Angiolillo, P., Yang, E. K. & Powers, L. Identification and assay of synchrotron radiation-induced alterations on metalloenzymes and proteins. *FEBS Lett.* **112**, 178–182 (1980).
2. Logan, D. T. *et al.* Crystal structure of reduced protein R2 of ribonucleotide reductase: the structural basis for oxygen activation at a dinuclear iron site. *Structure*. **4**, 1053–1064 (1996).
3. Harrenga, A. & Michel, H. The cytochrome c oxidase from *Paracoccus denitrificans* does not change the metal center ligation upon reduction. *J. Biol. Chem.* **274**, 33296–33299 (1999).
4. Schlichting, I. *et al.* The catalytic pathway of cytochrome p450cam at atomic resolution. *Science* **287**, 1615–1622 (2000).
5. Sjögren, T. & Hajdu, J. Structure of the bound dioxygen species in the cytochrome oxidase reaction of cytochrome cd1 nitrite reductase. *J. Biol. Chem.* **276**, 13072–13076 (2001).
6. Planche, L. A. Note sur la sophistication de la résine de jalap et sur les moyens de la reconnaître. *Bull. Pharmacie* **2**, 578–580 (1810).
7. Dunford, H. B. *Heme Peroxidases* (Wiley, New York, 1999).
8. Veitch, N. C. & Smith, A. T. *Advances in Inorganic Chemistry* Vol. 51, 107–162 (Academic, New York, 2001).
9. Weik, M. *et al.* Specific chemical and structural damage to proteins produced by synchrotron radiation. *Proc. Natl Acad. Sci. USA* **97**, 623–628 (2000).
10. Burmeister, W. P. Structural changes in a cryo-cooled protein crystal owing to radiation damage. *Acta Crystallogr. D* **56**, 328–341 (2000).
11. Leiros, H.-K. S., McSweeney, S. M. & Smalås, A. O. Atomic resolution structures of trypsin provide insight into structural radiation damage. *Acta Crystallogr. D* **57**, 488–497 (2001).
12. Lancaster, C. R., Kroger, A., Auer, M. & Michel, H. Structure of fumarate reductase from *Wolinella succinogenes* at 2.2 Å resolution. *Nature* **402**, 377–385 (1999).
13. Henderson, R. Cryoprotection of protein crystals against radiation-damage in electron and X-ray diffraction. *Proc. R. Soc. Lond. B*, **241**, 6–8 (1990).
14. Henderson, R. The potential and limitations of neutrons, electrons and X-rays for atomic resolution microscopy of unstained biological molecules. *Q. Rev. Biophys.* **28**, 171–193 (1995).
15. Gonzales, A. & Nave, C. Radiation damage in protein crystals at low temperature. *Acta Crystallogr. D* **50**, 874–877 (1994).
16. Neutze, R., Wouts, R., van der Spoel, D., Weckert, E. & Hajdu, J. Potential for biomolecular imaging with femtosecond X-ray pulses. *Nature* **406**, 752–757 (2000).
17. Ziaja, B., van der Spoel, D., Szöke, A. & Hajdu, J. Auger-electron cascades in diamond and amorphous carbon. *Phys. Rev. B* **64**, 214104-1–214104-8 (2001).
18. Persson, P., Lunell, S., Szöke, A., Ziaja, B. & Hajdu, J. Shake-up and shake-off excitations with associated electron losses in X-ray studies of proteins. *Protein Sci.* **10**, 2480–2484 (2001).
19. Nakajima, R. & Yamazaki, I. The mechanism of oxypoxidase formation from ferryl peroxidase and hydrogen peroxide. *J. Biol. Chem.* **262**, 2576–2581 (1987).
20. Van Wart, H. E. & Zimmer, J. Resonance Raman evidence for the activation of dioxygen in horseradish oxypoxidase. *J. Biol. Chem.* **260**, 8372–8377 (1985).
21. Hammes-Schiffer, S. Theoretical perspectives on proton-coupled electron transfer reactions. *Acc. Chem. Res.* **34**, 273–281 (2001).
22. Schlichting, I., Berendzen, J., Phillips, G. N. Jr & Sweet, R. M. Crystal structure of photolysed carbonmonoxy-myoglobin. *Nature* **371**, 808–812 (1994).
23. Edman, K. *et al.* High-resolution X-ray structure of an early intermediate in the bacteriorhodopsin photocycle. *Nature* **401**, 822–826 (1999).
24. Gajhede, M. *et al.* Crystal structure of horseradish peroxidase C at 2.15 angstrom resolution. *Nature Struct. Biol.* **4**, 1032–1038 (1997).
25. Chance, B. *et al.* X-ray absorption studies of intermediates in peroxidase activity. *Arch. Biochem. Biophys.* **235**, 596–611 (1984).
26. Szöke, A. Use of statistical information in X-ray crystallography with application to the holographic method. *Acta Crystallogr. A* **54**, 543–562 (2001).
27. Cravens, T. E. Comet Hyakutake X-ray source: Charge transfer of solar wind heavy ions. *Geophys. Res. Lett.* **24**, 105–108 (1997).

28. Henriksen, A., Smith, A. T. & Gajhede, M. The structures of the horseradish peroxidase C-ferulic acid complex and the ternary complex with cyanide suggest how peroxidases oxidize small phenolic substrates. *J. Biol. Chem.* **274**, 35005–35011 (1999).
29. Smith, A. T. *et al.* Expression of a synthetic gene for horseradish peroxidase C in *Escherichia coli* and folding and activation of the recombinant enzyme with calcium and heme. *J. Biol. Chem.* **265**, 13335–13343 (1990).
30. Read, R. J. Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallogr. A* **42**, 140–149 (1986).

Supplementary Information accompanies this paper on Nature's website (<http://www.nature.com>).

Acknowledgements

We thank M. Gajhede, K. G. Welinder, N. Veitch, B. Ziaja, D. Choudhury, M. Iwata, R. Subramanian and G. Katona for discussions and help. We thank MaxLab and ESRF/EMBO for beam time and services. This work was supported by the EU-Biotech Programme and by the Swedish Research Councils.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to J.H. (e-mail: janos@xray.bmc.uu.se).

retraction

Directed evolution of new catalytic activity using the α/β -barrel scaffold

M. M. Altamirano, J. M. Blackburn, C. Aguayo & A. R. Fersht

Nature **403**, 617–622 (2000).

This laboratory reported the *in vitro* evolution of an enzyme with phosphoribosyl anthranilate isomerase activity (ivePRAI) from the α/β -barrel scaffold of indole-3-glycerol-phosphate synthase using a combination of rationally designed libraries, DNA shuffling, and selection with *Escherichia coli*, JA300, a strain that lacks an active PRAI gene. As part of the ongoing project to characterize the structure and properties of ivePRAI, we discovered that the protein expressed from a variety of vectors that contained a synthetic gene corresponding to the sequence of ivePRAI as published is insoluble and does not complement JA300 (R. L. Weinberg, C. M. Blair and A.R.F., unpublished results), as reported. We conclude that the results are unsound.

It appears that the discrepancy in the results is due to a combination of two episodes of cross-contamination. We are now repeating the directed evolution of ivePRAI using modified procedures to test the design strategy that should eliminate the source of errors of contamination.

The first author of this Article (M.M.A.), who was responsible for most of the analysis and design strategy involving loop transfer and most of the experimental work, wishes to be dissociated from this retraction because she believes that the experimental data are fundamentally sound. □

Kit and kits for proteomics, genomics

Imaging, thermocycling, data mining, genotyping and cloning.

Proteomic Imager

PerkinElmer Life Sciences

<http://lifesciences.perkinelmer.com>

Multicolour fluorescence imaging

The ProXPRESS Proteomic Imaging System is designed to detect 0.5 ng to 1 µg of protein in a single SYPRO Ruby stained 2D gel. In addition, the ProXPRESS is claimed to save stained gels from the effects of bleaching. The system is optimized for multicolour differential protein detection in 2D gels and measures all fluorescent gel stains with consistent performance: SYPRO Ruby, Red, Orange and Tangerine, as well as silver and Coomassie Brilliant Blue stained gels. In addition, high-performance western blot imaging can be done using glow type chemiluminescence. Details of protocols for these techniques are available from the company web site.

Biotechnix 3d 1.1

Gentech

www.gentech.fr

A new dimension for Mac users

Gentech has released a version of Biotechnix 3d, its DNA and protein sequence analysis software package, for use with Apple Mac's newest operating system, OS X. The company comments that the power and stability provided by Mac OS X enhances the key feature of its software, three-dimensional visualization. This version is described as a flexible and productive set of tools for bioinformatics research.

DUALhybrid

Dualsystems Biotech

www.dualsystems.com

Get a rise out of this service

Dualsystems Biotech offers a customized, optimized dual hybrid screening service for *in vivo* identification of protein-protein interactions, based on the yeast two-hybrid system. DUALhybrid is said to incorporate improvements that allow rapid isolation of interacting protein pairs using cDNA library screens, with fewer false positives than occur with conventional yeast two-hybrid systems. Results are delivered within three months.

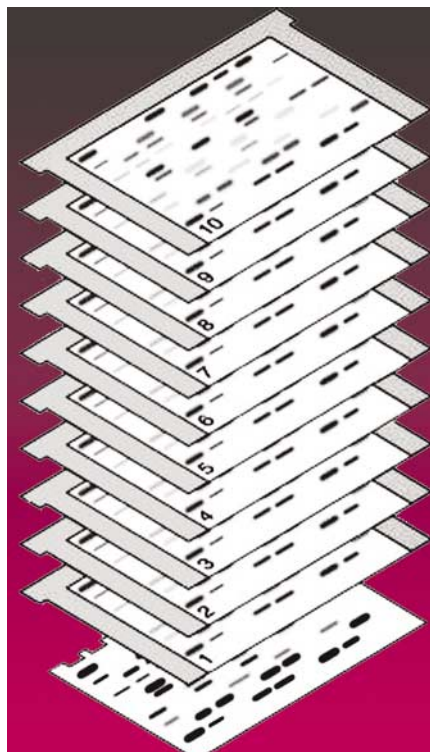
Bio-Max

Kodak

www.kodak.com

Multi-blot kit for proteins

Kodak's new kit allows the simultaneous production of 10 blots from a single protein gel. Compatible with existing western blotting tank transfer apparatus, the kit eliminates the need to run multiple gels or same-blot stripping and re-probing. Total protein distribution can be visualized (via membrane stain-



Kodak Bio-Max: ten blots from one gel.

ing) minutes after transfer, and each blot can be screened with one or more different antibodies.

Trobot

Biometra

www.biometra.com

Adopt me

The Trobot thermocycler is designed for integration into robotic systems. It features a motorized, heated lid, providing automated workflow and plate sealing. Once lifted, the plate is removed by a robotic arm. All features are controlled through a serial port. The Trobot can be used with existing software running individual robotic stations, and a program for operating up to eight Trobots



Trobot: run 8 of these from one PC.

from a computer is also available. The Trobot comes with either 96-well or 384-well blocks.

GenoLink

Bruker Daltonics

www.bdal.com

Speed up SNP genotyping

GenoLink is an automated accessory for MALDI-TOF mass spectrometry-based SNP genotyping. The device combines semi-automated sample preparation, automated MALDI-TOF analysis and bioinformatics to produce direct, label-free SNP genotyping. A robot performs semi-automated sample processing, magnetic bead DNA purification and MALDI target preparation. GenoLink can be linked with the company's MALDI-TOF mass spectrometers to detect SNPs within a few seconds per sample. By using GenoLink with the Proteineer suite, functional genomics researchers are able to analyse both genetic variations and protein expression.

GeneLinker Platinum

Molecular Mining Corporation

www.molecularmining.com

A grand slam, in silico

This data mining tool combines software and hardware to enable users to discover and characterize high-order correlations among objects in a multidimensional data set, using an algorithm known as SLAM (Sub Linear Association Mining). Objects may represent genes and toxicological response outcomes, or they may be chemical descriptors and binding assays. Together with a suite of clustering and visualization protocols, the system allows users to achieve large-scale, *in-silico* inference and prediction. Applications include SNP analysis, proteomics and gene expression.

ABI Prism GeneMapper

Applied Biosystems

www.appliedbiosystems.com

Software that pinpoints the problem

GeneMapper software for microsatellite and SNP-based genotyping applications uses a process component-based quality value system to minimize data examination efforts by filtering problematic data and pinpointing the origins of genotyping failure. Designed to support high-throughput genotyping projects, it will process 200,000 genotypes in 24 hours. Electropherogram traces are displayed in multiple colours with multiplexed marker and bin overlays. SNP marker labels appear immediately below peak clusters to enable quick review of allele calls in highly multiplexed samples.

SureScore

Invitrogen

www.invitrogen.com

Simple and accessible SNP detection

The SureScore genotyping kit from Invitrogen provides single nucleotide polymorphism detection in a simple format. It requires no specialized instrumentation and only minimal optimization, and is suitable for low- to medium-throughput studies.

Proteolink

Promega

www.promega.com

In vitro expression cloning system

This system uses a reticulocyte-based coupled expression system partnered with a plasmid-based human adult brain cDNA library to link phenotype (protein function) back to the genotype (gene). The product can be used for substrate identification, enzymatic activity and interaction screens. Advantages are said to include accurate protein folding, post-translational modifications and high-throughput screening capability.

BLAST

CAS

www.cas.org

One-stop sequence searching

Users of STN on the Web can now apply BLAST searching to the sequence information in the CAS Registry file and document records in STN databases. After identifying a nucleotide or protein sequence similar to a queried sequence, the user can view associated published research and patents in a range of bioscience databases and link out to the full text of the documents via ChemPort. The CAS Registry database contains over 18 million publicly available nucleotide and protein sequences derived from patents, journal literature and GenBank.

GeneGnome

Syngene

www.syngene.com

A mite more powerful

The GeneGnome for chemiluminescence detection and analysis provides rapid results with more than double the dynamic range of film. The image files are generated with a true 16-bit CCD. The system is GLP-compliant and can be expanded to work with gels that require ultraviolet or white light illumination for simultaneous acquisition. The compact unit measures 18 × 16 × 17 inches.

Pro-Pure protein assay

AMRESCO

www.amresco-inc.com

Copper-based protein determination

AMRESCO's Pro-Pure BCA Protein Determination Kit is a 1-litre kit based on a colorimetric test utilizing absorbance at 562 nm and a working range 20–2,000 µg per ml. The kit comprises BCA Solution (4 × 250 ml) and



Pro-Pure BCA, home in the 20–20,000 µg range.

copper solution (30 ml). This new kit joins AMRESCO's Pro-Pure line as a copper-based protein determination assay that is convenient to use, fast-acting, and completely compatible with detergents.

CEQ 8000

Beckman Coulter

www.beckmancoulter.com

Multitasking for genetic researchers

Beckman Coulter's CEQ 800 automates and integrates most genetic analysis functions into one flexible system. Using one gel, one array and one software platform, researchers can perform DNA sequencing, heterozygote detection, confirmatory sequencing, mutation analysis, allele identification, SNP scoring, microsatellite instability and AFLP fingerprinting. The capillary electrophoresis-based system is compact and provides both high-resolution and high-speed DNA analy-

sis. Its configuration of eight capillary arrays is suitable for use with 96-well plate formats.

Progenesis

Nonlinear Dynamics

www.nonlinear.com

Walk away from this one

Progenesis is a 'hands-free', automated, scaleable image analysis system designed to automate two-dimensional gel electrophoresis in high-throughput proteomics applications. It can be launched against high numbers of two-dimensional electrophoresis gel images for walk-away analysis that includes spot detection, measurement and matching.

PRMT antibodies

BioCarta

www.biocarta.com

A key to protein function

Arginine methylation is an irreversible post-translational modification, recently linked to protein activity. At least three types of protein arginine *N*-methyltransferase (PRMT) have been identified in mammalian cells. These enzymes have been shown to have essential regulatory functions by methylation of key proteins in several fundamental areas including nuclear proteins, cell cycle proteins, signal transduction proteins, apoptosis proteins, and viral proteins. BioCarta's new line of antibodies against human PRMTs can be utilized to explore arginine methylation and its regulatory functions.

These notes are compiled in the Nature office from information provided by the manufacturers.

Contacts

Publisher: Fabien Savenay

Editor: Paul Smaglik

Sales Director: Ben Crowe

European Head Office, London

The Macmillan Building
4 Crinan Street
London N1 9XW, UK
Tel +44 (0) 20 7843 4961
Fax +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

European Manager:

Nevin Bayoumi (4978)

UK/ RoW/ Ireland:

Matt Powell (4953)

Andy Douglas (4975)

Laura Pearson (4977)

Netherlands/ Italy/ Iberia:

Evelina Rubio Hakansson (4973)

Scandinavia: Silke Opstrup (4994)

Production Manager:

Billie Franklin
To send materials use London
address above.

Tel +44 (0) 20 7843 4814

Fax +44 (0) 20 7843 4996

e-mail: naturejobs@nature.com

International

Advertising Coordinator:

Hind Berrada (4935)

Naturejobs web development:

Tom Hancock

Naturejobs online production:

Ben Lund

European Satellite Offices

France/ Belgium:

Christine Niox-Chateau
Tel + 33 (0) 1 43 20 16 51
Fax + 33 (0) 1 43 20 51 52
e-mail: c.nioxchateau@nature.com

Germany/ Austria/ Switzerland:

Patrick Phelan/ Kate Turner
Tel + 49 89 54 90 57 11/-2
Fax + 49 89 54 90 57 20
e-mails: p.phelan@nature.com
k.turner@nature.com

US Head Office, New York

345 Park Avenue South,
10th Floor, New York, NY 10010-1707
Tel +1 800 989 7718
Fax +1 800 989 7103
e-mail: naturejobs@nature.com

US Sales Director:

Ben Crowe

US Sales Manager:

Peyton Mason

US Advertising Coordinator:

Ashly de Leon

Japan Head Office, Tokyo

MG Ichigaya Building (5F),
19-1 Harakata-machi,
Shinjuku-ku,
Tokyo 162-0841
Tel +81 3 3267 8751
Fax +81 3 3267 8746
e-mail: k.cowan@nature.jp
Japan Manager: Kate Cowan

naturejobs

Talent tug-of-war

The past few years have seen the United States increase its reliance on scientists and engineers from abroad, while at the same time, these workers' home countries have made a more determined effort to reclaim lost talent. This set of events could mean trouble for the United States, says the National Science Foundation's *Science and Engineering Indicators 2002* report, released this month (www.nsf.gov/sbe/srs/seind02).

According to the report, the percentage of foreign-born scientists and engineers in the United States is growing across all disciplines and at all degree levels. The proportions are at their highest for engineering (45%), computer sciences (43%) and mathematics (30%), all of which are fields that are currently showing little or no growth in the production of domestic PhDs.

If the ageing — and in some cases declining — pools of well-trained scientists in many industrialized nations are added to the equation, it is clear that there will be increased opportunities for young scientists outside the United States. A fact compounded by the efforts being made by many nations to build up their science and technology infrastructure and so recruit back their lost talent — a recurring theme in *Naturejobs'* Regions reports (see *Naturejobs* 4-5; 14 February 2002 and 4-5; 21 March 2002).

If this trend continues, the relative attractiveness of the United States as a destination for foreign-born scientists will decrease. "Although such a decline would be difficult to quantify," says the new report, "anecdotes suggest that experienced scientists and engineers, particularly those originally from Asia, are even now returning to their native countries."

So what does that mean for working scientists? More opportunities — especially if they work in maths, physics or chemistry and are willing to be flexible about where they live.

Paul Smaglik
Naturejobs editor



Contents

POSTDOCS & STUDENTS

Easing the transition into
postdoctoral research

p5

MOVERS

Wellcome Trust chief retires;
bioinformatics pioneer heads
for pastures new; French
research agency appoints life-
sciences head; and more

Back

WWW.NATUREJOBS.COM

Career centre
Information on the
scientific job market

FOCUS

SPOTLIGHT

RECRUITMENT

SCIENTIFIC ANNOUNCEMENTS

SCIENTIFIC EVENTS

Training programmes with vision

Wayne Streilein:
We're trying to break from the tradition that postdoctoral training must be an isolating experience.



Making the transition from PhD student to fully fledged research scientist is not always easy — and timing is often one of the problems. The trouble with many PhD students is that they do not think about their next step until they are about six months from graduating, says Thaddeus Dryja, director of the Cogan Eye Pathology Laboratory at the Massachusetts Eye and Ear Infirmary in Boston. But that leaves them too little time to write an individual postdoc proposal and get it approved without enduring a gap in funding, he says.

Fortunately, training grants from the US National Institutes of Health (NIH) that are awarded to departments rather than to individuals can ease this problem, smoothing the transition between PhD and postdoc projects.

So far, two postdocs who have passed through Dryja's lab began their studies on the molecular basis of eye diseases thanks to funding from a training grant. One of those postdocs, Stephanie Hagstrom, now an assistant professor of ophthalmology at the Cole Eye Institute at the Cleveland Clinic in Ohio, had applied for an individual grant from the NIH while finishing her PhD. But she, like many quality applicants, needed several more attempts before she met with success.

LOOKING GOOD

Through the training grant — the Training Program in the Molecular Basis of Eye Disease, which is shared among five eye-research institutes affiliated to Harvard University — Hagstrom got involved in a journal club, courses and poster sessions. It also allowed her to have regular meetings with Wayne Streilein, president of the Schepens Eye Research Institute in Boston, which administers the grant. "The programme encourages career planning and opened me up to more scientists and other postdocs," says Hagstrom.

Streilein believes that the training grants have started to redefine postdoctoral training, at least in his part of the research world. "We're trying to break from the tradition that postdoctoral training must be an isolating experience," he says.

Training grants were once relatively rare in ophthalmology. But the eye-disease programme has

led the way for all that to change. Until the late 1980s, the National Eye Institute (NEI) in Bethesda, Maryland, supported many postdoctoral training programmes. But, according to postdocs who received these fellowships, it used the money primarily to fund clinical fellows who did not do research.

In the early 1990s, Streilein and colleagues realized that the Boston–Cambridge area had a huge number of people working in vision research, but no formal programme to support trainees. So they appealed to the NEI to reconsider its position on postdoc training programmes.

"After two years of discussion, the NEI agreed on a pilot basis to let us train five trainees a year," says Streilein, who directs the Molecular Basis of Eye Disease programme. The award, which has been in existence for almost seven years, was renewed two years ago for another five years.

OPEN TO OTHERS

In light of the programme's success — and an element of relaxation in the NEI's postdoc funding rules — other organizations are now applying for similar programmes. Since 2000, the NEI has also allowed more than two postdoctoral positions to be included in all of its institutional training grants.

Another all-postdoc training programme is the Computational Visual Neuroscience Training programme at New York University's Center for Neural Sciences. Led by Robert Shapley, the programme is now nearing its first birthday. It was set up when Shapley successfully sought funding from the NEI to continue a programme seeded by the Alfred P. Sloan Foundation in New York a few years earlier.

The programme focuses on training postdocs from other fields such as mathematics and physics to work in vision research. Because this is a 'retooling' programme, the NEI allows for the possibility of more than one year of guaranteed funding for these postdocs, says Chyren Hunter, a training-programme officer at the NEI. The new flexibility on the part of the NEI is good, says Shapley, "in that it's bringing new people to vision research".

Karen Kreeger is a freelance science writer based in Philadelphia.

Web links

The Schepens Eye Research Institute
www.eri.harvard.edu
 NYU Center for Neural Science
www.cns.nyu.edu
 National Eye Institute
www.nei.nih.gov

Transitions

Three former Hewlett-Packard (HP) executives have joined the SETI Institute board of trustees: **Lewis Platt**, former HP chairman and chief executive; **Joel Bimbaum**, former HP senior-vice president of research and director of HP Labs; and **Alan Bagley**, another former HP executive. **Linda Bernardi**, founder and chief executive of ConnecTerra, has also joined the institute's board.

Protarga, a clinical-stage pharmaceutical company based in Pennsylvania, has appointed **Frank Bambarola** as vice-president, oncology, and **Timothy McKinlay** as senior director of clinical operations. Bambarola joins the company from Bristol-Myers Squibb, and McKinlay will be making his move from AstraZeneca.

Paul Lim, formerly clinical and regulatory director of AstraZeneca Singapore, was last month appointed to the Singapore office of Oxford Natural Products, where he will aid drug development. Oxford Natural Products is a biotech company that specializes in developing drugs from plant products.

This month, **Sylvain Goyon** will move from his current position as head of the European biotechnology team at CAI Chevreux to become director of corporate relations of NicOx, a pharmaceutical company based in Sophia Antipolis, France.

Robert Williamson, who was chief executive officer of the now-defunct genomics company DoubleTwist, has been hired as president and chief operating officer at Eos Biotechnology, a South San Francisco biotechnology company.

BIOINFORMATICS

In a switch from one Heidelberg company to another, bioinformatics pioneer **Georg Casari** has taken up the post of vice-president for informatics at biotech firm Cellzome. He joins the company from LION bioscience, which he co-founded in 1997. Both firms are spin-offs from the European Molecular Biology Laboratory (EMBL).

One of the reasons Casari made the move, he says, is because he enjoys trying to solve new kinds of bioinformatics challenges. LION now has its technology platform pretty much in place and is focusing more on drug discovery, he explains, whereas Cellzome is at an earlier stage of development. "The scientific challenge is much more at the core of the existence of the company," he says.

Casari is pleased that he will still stay near to EMBL in Heidelberg. He did his early scientific work there — including the development of high-throughput bioinformatics-system software, which enabled the then newly sequenced yeast genome to be annotated, as well as providing one of the seeds for LION's birth. Casari views companies as self-funding ways to advance science — in other words, they are alternatives to writing grants. "If you can make value for the customer, the money comes back for the development," he says.

DRUG DISCOVERY

Peter Fellner, already chief executive of one biotech company and a board member of another, last month added another position to his management portfolio. The Celltech chief executive is now non-executive chair of Astex Technology, a structure-based drug-discovery firm in Cambridge, UK. Fellner is also on the board of Oxford-based start-up Synaptica.

He acknowledges that his responsibilities — daily operation of Celltech plus giving regular strategic advice to Astex and attending board meetings of all three companies — will keep him busy. "One has to be quite organized for it to be done properly," he says. Fellner has years of biotech management experience, having been chief executive of Celltech for 10 years — "an unusually long time" by biotech industry standards, he says.

BIOTECHNOLOGY

In a career that has already taken him from the United States to Israel, **Stephen Runnels** is now heading for Singapore. He left his last job as chief executive of ProChon Biotech in Rehovot, Israel, in February, and will assume the same post at S*BIO later this summer. Partly sponsored by the Singaporean government, S*BIO is also partnered with Chiron and focuses on cancer research.



Georg Casari



Stephen Runnels



Peter Fellner



Roderic Pettigrew

CONTACTING US AT MOVERS

Please send any information on appointments or job moves to:
naturejobseditor@naturedc.com

Runnels is enthusiastic about the government's support for the company. "It is a very vibrant environment for the biomedical industry, with a great deal of interest in the success of the company coming from the economic development board," he says.

Apart from the usual drug-discovery and business-development challenges, Runnels is excited about establishing a mentorship programme that would place university students in the company's labs "so they can see there are careers in science outside academia".

FRENCH RESEARCH

The CNRS, France's main research agency, last month appointed **Bernard Pau** as director of its Department of Life Sciences, one of the agency's eight research divisions. Pau is currently professor of immunology and biotechnology at the University of Montpellier I and director of its Institute of Biotechnology and Pharmacology.

His career has taken him from industry to academia. Pau started out in the private sector, working first for the Coulter Group and then Sanofi, before taking up his university post in 1983. His work at the Institute of Biotechnology and Pharmacology served as a bridge between the two sectors.

BIOMEDICAL RESEARCH

Michael Dexter has announced his retirement as director of the Wellcome Trust, Britain's leading biomedical charity. He leaves his post next March, when his five-year term expires. Under Dexter's leadership, the \$15-billion foundation helped the Sanger Institute near Cambridge, UK, to grow into an international leader in genomics. At the same time, the trust was helping to shape British science policy, often by investing money for salary and infrastructure that the government was initially reluctant to spend. Dexter has not yet announced what his next project will be, and the trust has yet to select his successor.

BIOIMAGING

One year old, and the National Institutes of Health's youngest institute now has its first permanent director. **Roderic Pettigrew** of Emory University School of Medicine in Atlanta, Georgia, has been appointed to head the National Institute of Biomedical Imaging and Bioengineering, a post he will take up later this summer.

Pettigrew has spent much of his career developing cardiovascular imaging techniques, including dynamic three-dimensional imaging of the heart using magnetic resonance (MRI).